Egalitarianism under Incomplete Information*

Geoffroy de Clippel†

April 2010

Abstract

The paper aims at extending the egalitarian principle to environments with incomplete information. The approach is primarily axiomatic, focusing on the characteristic property of monotonicity: no member of the society should be worse off when more collective decisions are available. I start by showing the incompatibility of this property with incentive efficiency, even in quasi-linear environments. This serious impossibility result does not follow from the mere presence of incentive constraints, but instead from the fact that information is incomplete (asymmetric information at the time of making a decision). I then weaken the monotonicity property so as to require it only when starting from incentive compatible mechanisms at which interim utilities are transferable (in a weak sense). Adding other axioms in the spirit of Kalai's (*Econometrica*, 1977, Theorem 1) classical characterization of the egalitarian principle under complete information, I obtain a partial characterization of a natural extension of the lex-min solution to problems with incomplete information. Next, I prove that, in each social choice problem, there is a unique way of rescaling the participants' interim utilities so as to make this solution compatible with the ex-ante utilitarian principle. These two criteria coincides in the rescaled utilities exactly at the incentive efficient mechanisms that maximize Harsanyi and Selten's (*Management Science*, 1972) weighted Nash product. These concepts are illustrated on classical examples of profit-sharing, public good production and bilateral trade. The richness of the topic of social choice under incomplete information is illustrated by considering two alternative extensions of the egalitarian principle – one based on an idea of equity from the point of view of the individuals themselves (given their private information) instead of an uninformed third party (social planner or arbitrator), and another notion based on the idea of procedural justice.

*The paper beneficiated from stimulating discussions with David Perez-Castrillo and David Wettstein. I also wish to thank Kfir Eliaz for helpful comments on the exposition of the results. Financial support from the National Science Foundation (grant SES-0851210) is gratefully acknowledged.

†Department of Economics, Brown University. Email: declippel@brown.edu

1. INTRODUCTION

The theory of social choice has been applied extensively to determine collective actions. A limitation to its applicability, though, is the prevalent assumption of complete information. Indeed, in many practical scenarios, the participants already have some private information when making a collective decision. Developing models of cooperation under incomplete information has long been considered, and remains a significant open problem in economic theory, as pointed out, for instance, by Professor Aumann in his presidential address to the Game Theory Society (reproduced in Aumann, 2003). To be more precise, an impressive amount of work has already been devoted to understand which contracts are feasible under asymmetric information. Professors Hurwicz, Maskin, and Myerson were awarded the 2007 Sveriges riksbank prize in economic sciences in memory of Alfred Nobel for their path-breaking early contributions on the topic. Yet, little work has been done to identify contracts that are socially desirable among those that are feasible. Instead, researchers have devoted their effort to characterizing equilibrium outcomes of specific non-cooperative games (see e.g. the literatures on auctions, cheap talk, and contracts). I find that extending the theory of social choice to characterize selection criteria that would be applicable to the mechanism design problem is an important research agenda.

As a first step in that direction, the present paper discusses possible extensions of the egalitarian principle to environments with asymmetric information. The approach will be primarily axiomatic, focusing on the characteristic property of monotonicity, which is as straightforward under asymmetric information as it is under complete information: no member of the society should be worse off (whatever his or her private information) when more collective decisions are available.[1] In quasi-linear environments under complete information, the egalitarian solution – selecting the collective decision that maximizes the sum of the gains and designing monetary compensations so as to equalize those gains across individuals – is monotonic and systematically selects an outcome that is Pareto efficient. These two basic properties – Efficiency and Monotonicity – will be shown to become incompatible under incomplete information (Theorem 1). This impossibility result does not follow from the mere presence of incentive constraints, but instead from the fact that information is incomplete (asymmetric information at the time of selecting a collective decision). Indeed, no such impossibility would occur in quasi-linear environments if decisions about what to implement at the interim stage were made at the ex-ante stage, i.e. before the individuals learn their private information, or if asymmetric information involved moral hazard instead of adverse selection.

[1]Monotonicity properties of this type have a long tradition in the theories of social choice, distributive justice, and bargaining. First briefly discussed in the books of Luce and Raiffa (1957, pages 133 and 134) and Owen (1968), they have been systematically studied since then under the assumption of complete information, cf. Kalai and Smorodinsky (1975), Kalai (1977), Thomson and Myerson (1980), Kalai and Samet (1985), Young (1985a, 85b), Moulin and Thomson (1988), and Moulin (1992), to cite only a few references that illustrate the variety of environments where it has been studied. Monotonicity is also necessary to provide individuals with incentives to search for new profitable collective decisions, and to divulge these new options in all circumstances.

Efficiency and Monotonicity are also incompatible under complete information in the absence of "transferable utilities" (this is known at least since Luce and Raiffa (1957)). It is helpful to understand why. Consider a first problem where only one collective decision, d, is available to replace the status-quo, and assume further that the first participant is indifferent between the alternative and the status-quo, while the second strictly prefers the former. Efficiency will guide the social planner or the arbitrator in selecting the alternative over the status-quo. Consider then a second problem that is symmetric to the first, in that only one collective decision, d', is available to replace the status-quo, with the property that the second participant is indifferent between the alternative and the status-quo, while the first strictly prefers the former. Again, efficiency will guide the social planner or the arbitrator in selecting the alternative over the status-quo. Let then the third problem be the union of the first two: any lottery selecting either d or d' is feasible. Clearly, there is no way to solve uniquely that new problem so as to make both player 1 better off than with the solution to the second problem, and player 2 better off than with the solution to the first problem. It has long been understood that this kind of impossibility is due to an extreme lack of utility transferability (e.g. starting from d, there is no alternative decision in the first problem that would allow to make player 1 strictly better off, even if one is ready to make the second individual worse off in any amount). A usual way to deal with this complication is actually to avoid it altogether by restricting the class of acceptable social choice problems. A classical example is Moulin's (1988, Theorem 3.2) textbook presentation of Kalai's (1977, Theorem 1) now classical axiomatic characterization of the egalitarian solution,[2] where social choice problems are assumed to satisfy a property of "minimal transferability,"[3] meaning that, at any feasible contract that is efficient and individually rational, and for any individual i, there exists an alternative feasible option that makes all the other individuals better off (thereby at the expense of i).[4] In particular, it rules out the possibility of satiation.

Theorem 1 thus also shows that restricting the class of social choice problems is not anymore a practical way to resolve the incompatibility of Efficiency and Monotonicity under incomplete information, as this incompatibility already occurs on the very restrictive class of quasi-linear social choice problems (i.e. with unlimited one-to-one transferability ex-post). The reason is that incentive constraints may lead to feasible sets of interim utilities that are non-comprehensive (for instance, a type of an individual may beneficiate of an "informational rent" in any incentive efficient mechanism) and with the possibility of satiation, even in the simplest quasi-linear environments. The complexity of how incentive constraints may influence the shape of the feasible sets of utilities at the interim

[2]To be precise, Kalai actually characterized the proportional solutions in his original paper, but of course the egalitarian solution is the only one to be anonymous in that class.

[3]Kalai's original result only required free disposal, which is weaker than the property of minimal transferability, but at the cost of accommodating only a weak form of Pareto efficiency, which is not very appealing, especially in social choice theory.

[4]Though not explicit in its name, the usual notion of transferable utility in cooperative games (or of quasi-linearity in the previous paragraph) requires in addition that, starting from *any* contract, utilities can be transferred at a *constant* rate of *one to one*. Minimal transferability is thus far weaker, to a point that it is almost unrelated.

stage is a difficulty that will be present throughout the paper, implying that axiomatic results are definitely more challenging to derive than under complete information.

As a variant to the domain restriction, Theorem 1's impossibility will be avoided by weakening the monotonicity property in a similar spirit. I will say that interim utilities are[5] "transferable" at an incentive compatible mechanism μ if for any two participants i, j, and any possible pair of types, t_i for i and t_j for j, there exists an alternative incentive compatible mechanism ν that is better than μ for i when of type t_i, worse than μ for j of type t_j, and at least as good as μ for any combination of participant and type that is different from (i, t_i) and (j, t_j). The axiom of "Restricted Monotonicity" (R-MON) is then the monotonicity property applied only when starting from a mechanism at which interim utilities are transferable. A social planner or an arbitrator may feel very constrained at a mechanism where interim utilities are not transferable because he would like to pick an alternative mechanism that is more favorable to individual i, when of type t_i, at the expense of individual j, when of type t_j, but cannot do so because of the incentive and the feasibility constraints. In such cases, having more collective decisions available may soften this constraint, and result in a mechanism that is less favorable to individual j of type t_j. I am not claiming that monotonicity should systematically be violated when adding collective decisions to a problem whose solution is a mechanism at which interim utilities are not transferable. Instead I am arguing that these are cases where monotonicity *might* be problematic and less appealing. R-MON thus remains silent in those dubious cases.

R-MON is compatible with Holmström and Myerson's (1983) notion of incentive efficiency (I-EFF). In Section 5, I will introduce additional axioms in the spirit of Kalai (1977, Theorem 1) in order to extend his characterization of the egalitarian principle to problems with incomplete information. Anonymity requires that renaming individuals and/or their types should have a covariant impact on the solution of any problem. Most of the theory of social choice under complete information (including Kalai's result, of course) is phrased under the welfarist assumption that only feasible utilities matter, not the underlying decisions that make them feasible.[6] Though it is certainly worthwile to find more primitive properties to justify that assumption,[7] or to study context-dependent social choice functions that violate it,[8] I feel that it is more natural to start by trying to extend the most standard approach to frameworks under incomplete information before finding interesting ways of departing from the benchmark. Understanding what is the right notion of welfarism under incomplete information is not that obvious in itself. A first idea that may come to mind is to require that only the sets of utility vectors that are feasible ex-post (i.e. one set for each possible realization of the types), should be sufficient information to determine the solution. This approach is necessarily wrong, as

[5]Again, transferability here should not be confused with its narrow meaning in cooperative games under complete information - cf. footnote 4.

[6]This welfarist assumption that remains implicit in the way classical models are phrased was first emphasized by Roemer (1986).

[7]In the context of social choice under complete information, see e.g. Moulin (1985), Roemer (1986), Gines and Marhuenda (2000), and de Clippel and Bejan (2009).

[8]Cf., for instance, the concepts of envy-freeness (Foley (1967), Varian (1974), and Feldman and Kirman (1974)) and egalitarian equivalence (Pazner and Schmeidler (1978)).

it does not allow to keep track of incentive constraints. Indeed, the fact that a utility vector is feasible at some type profile does not allow to infer what would be the utility that an individual would get should he report a different type. Also, only interim preferences matter when taking a decision alone under incomplete information, and therefore one may take the position that interim data should be sufficient as well when taking a collective decision. This motivates the addition of an axiom of "Interim Welfarism" (I-WELF): if the set of interim utilities that can be generated via incentive compatible mechanisms is the same in two social choice problems, then the set of interim utilities that can be generated by the mechanisms in the solution should be the same for both problems. Also, it is often understood in a welfarist model that any feasible option that leads to the same utilities as an element in the solution should be a valid choice as well. This property will now be imposed explicitly under the name of "Exhaustivity" (EX). Theorem 2 then shows that these axioms – I-EFF, R-MON, AN, I-WELF, and EX – offer a partial characterization of a natural extension of the egalitarian principle[9] to problems with incomplete information, which I call the *interim lex-min solution*. Given a social choice problem, one may compute the interim utility for every type of every individual associated to any given incentive compatible mechanism. After ranking these utilities increasingly, the interim lex-min solution then picks the mechanism that maximizes these vectors of interim utilities according to the lexicographic ordering. Theorem 2' provides a variant of Theorem 2 that proves more useful in some applications. Details are deferred to the main text.

Interpersonal comparisons of interim utilities come as a consequence of the axioms, of course as it does under complete information. I react in three ways to this fact. First I apply the new criterion to quasi-linear examples (sharing the profit of a collectively-owned technology, sharing the cost of production of a public good, and determining the fair price in a bilateral trade problem). Interpersonal comparisons are indeed easiest to accept in the presence of a numeraire. Many examples in mechanism design already fall in that category, because quasi-linearity simplifies a bit the often-difficult task of characterizing incentive efficient mechanisms. Second, I pursue Harsanyi's (1963) methodology (see also Shapley (1969) and Yaari (1981)) of endogenizing interpersonal comparisons so as to reconcile the utilitarian and the egalitarian principles. Here I aim at combining the ex-ante utilitarian criterion (which is often considered informally as a natural extension of the utilitarian criterion to the interim stage; see Nehring (2004) for a first formal argument in that direction) and my interim egalitarian criterion by rescaling the interim utilities. Interestingly, it turns out (see Theorem 2) that this is always feasible, even while requiring the interim egalitarian criterion to hold with equality (no need to resort to the lex-min refinement), it leads to a unique solution, and results in a characterization of Myerson's

[9] The basic meaning of egalitarianism under complete information is often understood as choosing an option that is Pareto efficient and that equalizes the individuals' utility gains. Yet, it is obvious that such options are not always available - cf. the description of the equality/efficiency dilemma in Moulin (1988, Chapter 1). Actually, it is precisely in those cases that the property of monotonicity may be violated. The usual adaptation of the basic principle, so as to apply in all problems, is to maximize a lex-min ordering. The interim lex-min is thus a natural extension of that workable version of the egalitarian principle to frameworks involving incomplete information.

(1979) solution, maximizing Harsanyi and Selten's (1972) weighted Nash product over the set of interim utilities that are achievable through some incentive compatible mechanism (see Weidner (1992), for a direct axiomatic characterization of that solution under the assumption of independent types). Third, in a companion paper (de Clippel et al., 2010), the solution is applied to utility gains that are measured endogenously in a way that extends Pazner and Schmeidler's (1978) concept of egalitarian equivalence to economies under asymmetric information. Though we do not have an axiomatic characterization of that variant, it has the advantage of being ordinally invariant, and applicable to rational preference orderings that do not necessarily satisfy the expected utility hypothesis.

I hope that the present paper will serve as a benchmark for future work in social choice under incomplete information. I see the interim lex-min solution as the most straightforward extension of the egalitarian principle. While it does capture obvious equity considerations – viz. the uninformed third party (social planner or arbitrator) trying to make sure that all the individuals enjoy the same expected benefits given their own private information, and this whatever the actual profile of types – I believe that there might be alternative appealing ways to proceed. To illustrate the potential richness of the topic of social choice under incomplete information, I suggest in the concluding section two other routes to derive natural extensions of the egalitarian principle, and show by means of examples how they differ in their prescriptions. In the first approach, I propose to place a higher weight on the individuals' knowledge at the interim stage to determine what is equitable. Indeed, it will become clear after studying some examples that, even though mechanisms in the interim lex-min solution are equitable from the point of view of an uninformed third party, some individuals may feel, *given the private information they happen to have*, that the outcome of these mechanisms is actually biased in favor of some other individuals. This motivates another possible extension of egalitarianism in quasi-linear collective choice problems that selects the incentive efficient mechanisms that maximize the minimum of the type-agents' ratios between their expected utility gains and the total surplus they expect the mechanism to realize. In a second approach, I discuss briefly the procedural approach to equity and distributive justice. We will see that the equivalence between the egalitarian principle and the equilibrium outcome of the random dictatorship procedure in quasi-linear problems with risk-neutral individuals does not survive the presence of asymmetric information.

The rest of the paper is organized as follows. The related literature is discussed in the next section. Section 3 defines the general model and notations. The incompatibility of I-EFF and MON in quasi-linear problems is presented in Section 4. I introduce the interim lex-min solution in Section 5, and present the two partial axiomatic characterizations involving R-MON and I-WELF. In Section 6, I show how there is a unique way to rescale the individuals' interim utilities so as to make the ex-ante utilitarian principle compatible with my interim egalitarian principle, leading to a new characterization of Myerson's (1979) variant of Harsanyi and Selten's (1972) solution. Some applications in quasi-linear environments are analyzed in Section 7. Alternative criteria and subjects of reflection for further research in social choice under incomplete information are presented in Section 8.

2. RELATION TO THE LITERATURE

I am aware of very few papers directly studying social choice criteria under incomplete information from an axiomatic perspective. Nehring (2004) studies interim social welfare orderings that allow to compare profiles of ex-post utilities. He shows that two axioms of consistency, one with the interim Pareto criterion and the other with uniform ex-post utilitarian comparisons, are compatible if and only if the individuals' beliefs can be derived from a common prior. In addition, the only interim social welfare ordering that satisfies these two axioms when a common prior exists, is the ex-ante utilitarian criterion. I show in de Clippel (2010) that Nehring's methodology – extending to the interim stage classical social welfare orderings by combining ex-post arguments with the interim Pareto criterion – essentially works only for the utilitarian criterion. Indeed, his two axioms become incompatible for most common priors once the utilitarian criterion is replaced by any other social welfare ordering that satisfies the strict Pigou-Dalton transfer principle (as does the egalitarian principle, for instance). As a consequence, and given that the enforceable collective decision is made at the interim stage, I drop the idea of consistency with ex-post comparisons, and instead try to find the criterion that emerges from a direct adaptation of axioms that are known to capture the essence of the egalitarian principle under complete information. Also, I study social welfare functions instead of social welfare orderings, and I take into account the incentive constraints explicitly. Olszewski (2004) characterized Moulin's (1994) serial cost sharing method for the provision and pricing of excludable public goods. He proves that the serial method is Pareto dominant within the class of mechanisms that are continuous, individually rational, coalition strategy-proof, and treats equals equally. Maniquet and Sprumont (2010) also show that the serial method is the only second-best method (within the class of strategy-proof mechanisms) that satisfies an axiom of demand monotonicity (assuming that there is a large population, that the cost function is convex, and that the quasi-linear preferences satisfy the single-crossing property). This class of results, and the rule they characterize, is specific to the problem of provision and pricing of excludable public goods. The approach I propose in the present paper is applicable to any social choice problem. Of course, as it is the case under complete information, this comes at the cost that the analysis cannot be "context-dependent." In particular, the serial method cannot be phrased in the space of utilities, and hence is ruled out by the axiom of interim welfarism. Another difference in methodology is the fact that my arguments are phrased with Bayesian incentive compatibility constraints instead of strategy-proofness.

There is a small literature on axiomatic bargaining under asymmetric information, which might be relevant since there is a strong link under complete information between the theories of bargaining and social choice, the Nash product, for instance, being a focal criterion in both fields. The first extension of Nash's bargaining theory is due to Harsanyi and Selten (1972), who start by discussing which profiles of interim utilities are achievable via demand games, and then, abstracting from the strategies that lead to these profiles of interim utilities, characterize a unique solution which amounts to maximize a weighted Nash product, where types are weighted according to their marginal likelihood. Myerson (1979) then argues that their criterion should be maximized over a larger set of interim utilities, namely those that can be achieved via incentive compatible mechanisms, as a consequence of the revelation principle. Weidner (1992) reconstructs Harsanyi and Selten's characterization result of the weighted Nash product on Myerson's domain, and argues that their result then holds only when the players' types are independently distributed. Myerson (1984a) takes a radically different standpoint, arguing against the Harsanyi-Selten solution for bargaining problems, independently of the set of interim utilities over which it is maximized. One of his arguments is that the solution is sensitive to joint changes of the utility functions and the players' beliefs that nevertheless leave interim preferences unchanged – formally a violation of his probability-invariance axiom. So indeed the Harsanyi-Selten solution is not well-defined if it is to be applied to Savage-type preferences, where beliefs are subjective. This criticism would disappear, though, in the more restrictive case where uncertainty is objective. Although I have not investigated this idea carefully from an epistemic perspective, I think that the probabilities could also be interpreted in the social choice perspective developed in the present paper as the subjective probability of the uninformed benevolent third party. Notice though that the interim lex-min solution is not subject to Myerson's criticism, as it does satisfy his probability invariance axiom (which is implied by my axiom of interim welfarism). So, in view of Theorem 2, the fact that the Harsanyi-Selten solution violates Myerson's probability invariance axiom can be traced back to the fact that the ex-ante utilitarian principle violates it. Another criticism from Myerson (1984a) is that the Harsanyi-Selten solution does not coincide with the equilibrium outcome of a simple bargaining procedure when it seems that any bargaining solution should – formally a violation of his random dictatorship axiom. This argument from non-cooperative behavior makes a lot of sense in bargaining, especially in the context of the Nash program, but less in social choice. Even so, I will go back to Myerson's idea in Section 8 to show how a procedural approach to justice might lead to recommendations that differ from those made by the interim lex-min solution, or even the variant briefly discussed at the end of the Introduction, a feature that is due solely to the presence of asymmetric information. As already mentioned earlier, characterization results are significantly more difficult to establish under incomplete information than under complete information because one needs to keep track of incentive constraints and understand how ex-post utilities and beliefs influence the shape of the feasible sets of interim utilities. The virtual utility construction of Myerson's (1984a) analysis of the bargaining problem will prove very useful in the present paper as well (the Appendix contains a refresher).

Finally, there exists a slightly more extensive literature whose objective is to compute the optimal mechanism under a given social choice criterion, without discussing the normative appeal of the criterion itself. Mirrlees (1971) is a first classic example in that category. The social objective he follows is to maximize the sum of the agents' utilities (or a common transformation of those utilities), in the utilitarian tradition. His methodology has been followed since then in the literature on optimal taxation. In most papers, there is a large population, which implies that all possible types (representing, for instance, the agents' productivity or their cost of effort) are present in the population. Classic social welfare criteria that were defined under complete information can thus be applied successfully, since they all treat individuals anonymously anyway. Another classic paper where a utilitarian principle is applied to select an incentive compatible mechanism is Myerson and Satterthwaite's (1983) bilateral trade problem. Here, only two agents are interacting and only one pair of types (interpreted as reservation prices) is actually realized. The utilitarian criterion is applied ex-ante, i.e. behind the veil of ignorance and using the relative likelihood of each possible pair of types. To the best of my knowledge, Nehring's (2004) result that I discussed earlier is the only existing axiomatic characterization of that ex-ante utilitarian criterion. There is a more recent literature that is developing at the intersection of computer science and economics that looks for strategy-proof mechanisms that maximize a worst-case scenario index, in order to guarantee, for instance, a minimal percentage of the maximal total surplus in every possible realization of the types (see e.g. Guo and Conitzer, 2009, Moulin, 2009, and references therein). Though intuitively appealing, this criterion has not been axiomatically characterized yet. Another difference, of course, is that their approach is non-Bayesian. More importantly, we see that all these contributions focus on the utilitarian principle. The present paper not only offers an axiomatic discussion of a normative criterion that is applicable in numerous practical problems, but also suggests an alternative point of view of distributive justice under incomplete information, focusing on equity instead of some form of sum-efficiency.

3. GENERAL MODEL

A *social choice problem under incomplete information* is a sextuple

$$\mathcal{S} = (I, D, d^*, (T_i)_{i \in I}, p, (u_i)_{i \in I}),$$

where I is the finite set of *individuals*, D is the set of *collective decisions*, $d^* \in D$ is the *status-quo*, T_i is the finite set of individual i's possible *types*, $p \in \Delta(T)$ is the *common prior* determining the individuals' *beliefs* (where $T = \times_{i \in I} T_i$), and $u_i : D \times T \to \mathbb{R}$ is individual i's *utility function*, that will be used to determine his interim preferences via the expected utility criterion. It will be assumed for notational convenience that $u_i(d^*, t) = 0$, for all $t \in T$. This is without loss of generality if utilities are understood as representing gains over the status-quo. I will also assume that T is the only nonempty common knowledge event. This is without loss of generality, as the results can be applied over minimal common knowledge events, and then merged so as to apply to the whole set of type profiles.

A (direct) *mechanism* for $\mathcal{S}$ is a function $\mu : T \to \Delta(D)$. If a mechanism μ is implemented truthfully, then individual i's expected utility when of type t_i is given by:

$$U_i(\mu|t_i) = \sum_{t_{-i} \in T_{-i}} p(t_{-i}|t_i) u_i(\mu(t), t).$$

If all the other individuals report their true type, while individual i reports t'_i instead of his true type t_i, then his expected utility is denoted $U_i(\mu, t'_i|t_i)$:

$$U_i(\mu, t'_i|t_i) = \sum_{t_{-i} \in T_{-i}} p(t_{-i}|t_i) u_i(\mu(t'_i, t_{-i}), t).$$

The mechanism μ is *incentive compatible* if

$$U_i(\mu|t_i) \geq U_i(\mu, t_i'|t_i)$$

for each t_i, t_i' in T_i and each $i \in I$. The revelation principle (Myerson, 1979) implies that any agreement that is achievable through some form of communication can also be achieved through truth-telling in an incentive compatible direct mechanism. Hence one may restrict attention to those mechanisms without loss of generality.

An incentive compatible mechanism μ is *interim individually rational* if $U_i(\mu|t_i) \geq 0$, for all $t_i \in T_i$ and all $i \in I$. A mechanism is *feasible* if it is both incentive compatible and interim individually rational. The set of feasible mechanisms will be denoted $\mathcal{F}(\mathcal{S})$.

Let $\mathcal{U}(\mathcal{M})$ be the set of interim utilities that can be achieved via incentive compatible mechanisms that belong to a set $\mathcal{M}$:

$$\mathcal{U}(\mathcal{M}) = \{\mathfrak{u}(\mu) | \mu \in \mathcal{M}\},$$

where $\mathfrak{u}(\mu) = (U_i(\mu|t_i))_{t_i \in T_i, i \in I}$. It is easy to check that $\mathcal{U}(\mathcal{F}(\mathcal{S}))$ is a convex set, as first observed by Myerson (1979). For notational simplicity, I will restrict attention to social choice problems for which there exists $\mathfrak{u} \in \mathcal{U}(\mathcal{F}(\mathcal{S}))$ such that $\mathfrak{u} >> 0$, and for which $\mathcal{U}(\mathcal{F}(\mathcal{S}))$ is compact. This last assumption is true whenever D is finite and in usual applications that involve a continuum of collective decisions, but it is possible to construct mathematical examples that would violate it.

A *social choice function* is a correpondence Σ that associates a nonempty set of feasible mechanisms to each social choice problem: $\Sigma(\mathcal{S}) \subseteq \mathcal{F}(\mathcal{S})$, for each $\mathcal{S}$. Even though correspondences are allowed, it is assumed that a social choice function is essentially single-valued, in the sense that all the individuals must be indifferent (whatever their private information) between any two mechanisms that belong to the solution of any problem $\mathcal{S} = (I, D, d^*(T_i)_{i \in I}, p, (u_i)_{i \in I})$:

$$(\forall \mu, \mu' \in \Sigma(\mathcal{S}))(\forall i \in I)(\forall t_i \in T_i) : U_i(\mu|t_i) = U_i(\mu'|t_i). \tag{1}$$

For reference, a social choice problem $\mathcal{S} = (I, D, d^*, (T_i)_{i \in I}, p, (u_i)_{i \in I})$ is *quasi-linear* if there exist a set $\hat{D}$ and functions $\hat{u}_i : \hat{D} \times T \to \mathbb{R}$ (one for each $i \in I$) such that $D = \hat{D} \times \mathbb{R}^I$ and $u_i((\hat{d}, m), t) = \hat{u}_i(\hat{d}, t) + m_i$, for all $(\hat{d}, m) \in \hat{D} \times \mathbb{R}^I$.

4. IMPOSSIBILITY

The purpose of this section is to establish the strong incompatibility, under incomplete information, of the properties of incentive efficiency (Holmström and Myerson (1983)) and monotonicity.

Incentive Efficiency (I-EFF) *Let $\mathcal{S}$ be a social choice problem, and let $\mu \in \Sigma(\mathcal{S})$. Then there does not exist a mechanism $\hat{\mu} \in \mathcal{F}(\mathcal{S})$ such that $U_i(\hat{\mu}|t_i) \geq U_i(\mu|t_i)$, for each $t_i \in T_i$, and each $i \in I$, with at least one of the inequalities being strict.*

Monotonicity (MON) *Let $\mathcal{S}$ and $\mathcal{S}'$ be two social choice problem. Suppose that $\mathcal{S}'$ differs* *from $\mathcal{S}$ only in that more collective decisions are available: $I = I'$, $D \subseteq D'$, $T_i = T'_i$, and $u_i(d,t) = u'_i(d,t)$, for each $i \in I$, each $d \in D$, and each $t \in T$. If $\mu \in \Sigma(\mathcal{S})$ and $\mu' \in \Sigma(\mathcal{S}')$, then $U_i(\mu'|t_i) \geq U_i(\mu|t_i)$, for each $t_i \in T_i$, and each $i \in I$.*

I-EFF requires the social choice function to systematically exhaust the benefit of cooperation at the time of agreeing. R-MON means that no individual in the society should be worse off when more collective decisions become available.

Theorem 1 *There is no social choice function that satisfies both I-EFF and MON, even on the restricted class of quasi-linear social choice problems.*[10]

Proof: The proof goes by way of example. Consider a social choice problem with two individuals,[11] 1 and 2, that can be of two types, L or H. Each individual knowns only his own type, and believes that the two types of the other individual are equally likely. Each individual has up to 10 hours available to work, and his productivity per hour is 1 if his type is L, and 2 if his type is H. Allowing for any kind of transfers and free disposal, the set of decisions is thus

$$D = \{(\alpha_1, \alpha_2, m_1, m_2) \in [0,10]^2 \times \mathbb{R}^2 | m_1 + m_2 \leq 0\}$$

and the utility functions are given by the following expression:

$$u_i((\alpha,m),t) = \pi_i(t_i)\alpha_i + m_i,$$

for each $(\alpha,m) \in D$, each $i \in \{1,2\}$ and $t_i \in \{L,H\}$, with the convention $\pi_i(L) = 1$ and $\pi_i(H) = 2$, for each $i \in \{1,2\}$. One may think of each individual cultivating a similar field, their payoffs being the quantity produced on their own field, which depends on their productivity, modified by any kind of monetary subsidy and taxes. We will assume that $d^* = 0$.

Let's consider now a feasible mechanism[12] (α, m) that determines a decision in D as a function of the individuals' reports.[13] The incentive constraints faced by the first individual can be written as follows:

$$\bar{m}_1(H) - \bar{m}_1(L) \leq \bar{\alpha}_1(L) - \bar{\alpha}_1(H) \leq \frac{\bar{m}_1(H) - \bar{m}_1(L)}{2} \tag{2}$$

where $\bar{\alpha}_1(L)$ (resp. $\bar{\alpha}_1(H)$) is the average quantity of time the first individual thinks he will have to work given the mechanism when of type L (resp. H), i.e.

$$\bar{\alpha}_1(L) = \frac{1}{2}(\alpha_1(L,L) + \alpha_1(L,H)) \text{ and } \bar{\alpha}_1(H) = \frac{1}{2}(\alpha_1(H,L) + \alpha_1(H,H)),$$

[10]The larger the domain, the easier it is to prove an impossibility result. Also, as indicated in the Introduction, the result is not surprising if one includes social choice problems with a severe lack of utility transferability. The unexpected character of Theorem 1 is that the impossibility may occur even in problems with unlimited one-to-one transferability.

[11]The example extends in a straightforward way to accomodate any number of individuals, but at the cost of heavier notations. The qualitative results remain unchanged.

[12]To keep notations lighter, the same letters are used to denote a collective decision, and the mechanism that selects a collective decisions as a function of the type reports.

[13]Utilities being linear in both time and money, there is no loss of generality in discussing only deterministic mechanisms.

and $\bar{m}_1(L)$ (resp. $\bar{m}_1(H)$) is the average quantity of money the first individual thinks he will receive given the mechanism when of type L (resp. H), i.e.

$$\bar{m}_1(L) = \frac{1}{2}(m_1(L,L) + m_1(L,H)) \text{ and } \bar{m}_1(H) = \frac{1}{2}(m_1(H,L) + m_1(H,H)).$$

Equation (2) implies that $\bar{m}_1(H) \leq \bar{m}_1(L)$ and $\bar{\alpha}_1(L) \leq \bar{\alpha}_1(H)$. If the mechanism is incentive efficient, then it must be that $\bar{\alpha}_1(H) = 10$. Otherwise, one could construct another feasible mechanism that interim Pareto dominates (α, m) by slightly increasing both $\bar{\alpha}_1(L)$ and $\bar{\alpha}_1(H)$ by a same amount, while keeping α_2 and m unchanged. Notice also that the second inequality in (2) must be binding if (α, m) is incentive efficient. Indeed, suppose on the contrary that the inequality is strict. Hence $\bar{\alpha}_1(L) < 10$ (as otherwise $\bar{\alpha}_1(L) = \bar{\alpha}_1(H)$, and (2) implies that $\bar{m}_1(L) = \bar{m}_1(H)$, which contradicts the fact that the second inequality is strict). Now one can construct another feasible mechanism that interim Pareto dominates (α, m) by increasing a bit $\bar{\alpha}_1(L)$, while keeping $\bar{\alpha}_1(H)$, α_2 and m unchanged. This contradicts the fact that (α, m) is incentive efficient, and thereby proves that (2) is binding. Notice now that $\bar{\alpha}_1(L)$ must equal 10, as well. Otherwise, consider an alternative mechanism where $\bar{\alpha}_1(L)$ is increased by a small amount ϵ, while keeping $\bar{\alpha}_1(H)$ and $\bar{\alpha}_2$ constant, and changing monetary transfers as follows: $\Delta m(L,L) = \Delta m(L,H) = (-\epsilon, +\epsilon)$ and $\Delta m(H,L) = \Delta m(H,H) = (+\epsilon, -\epsilon)$. Since $\bar{\alpha}_1(L) < 10$ and the second inequality of (2) is binding, it must be that the first inequality of (2) is strict for the original mechanism. The change makes lying for player 1 a bit more attractive when of a low type, but not enough for him to actually lie if ϵ is small enough. The incentive constraint remains binding when he is of a high type. As for player 2, nothing changes for him, since he ignores player 1's type, and the additional tax of ϵ when player 1 reports a high type is exactly compensated by the additional subsidy of ϵ when player 1 reports a low type. In terms of interim payoffs, both types of both individuals get at least as much with the new mechanism than with the original one, but player 1 gets strictly more when of a high type, thereby contradicting the fact that the original mechanism is incentive efficient. Hence $\bar{\alpha}_1(L) = \bar{\alpha}_1(H) = 10$, and (2) implies $\bar{m}_1(L) = \bar{m}_1(H)$. A similar reasoning applies to individual 2. Hence, if a mechanism (α, m) is incentive efficient, then there exists $(x_1, x_2) \in \mathbb{R}^2$ such that $x_1 + x_2 = 0$, and

$$U_i((\alpha, m)|L) = 10 + x_i \text{ and } U_i((\alpha, m)|H) = 20 + x_i,$$

for $i = 1, 2$. Conversely, any such interim payoffs can be achieved by an incentive compatible mechanism (α, m), where $\alpha_i(t) = 10$, $m_i(t) = x_i$, for each $t \in \{L, H\}^2$ and each $i \in \{1, 2\}$.

Consider now a similar problem, but where the two individuals work on a third field, in which case a total productivity of 3 per joint hour of work can be achieved. They can decide in which proportion the output will be shared, and to implement monetary compensations. Formally, the set of collective decisions is

$$D' = \{(\alpha'_1, \alpha'_2, s_1, s_2, m_1, m_2) \in [0, 10]^2 \times [0, 1]^2 \times \mathbb{R}^2 | s_1 + s_2 = 1 \text{ and } m_1 + m_2 \leq 0\}$$

and the utility functions are given by the following expression:

$$u'_i((\alpha', s, m), t) = 3s_i \min\{\alpha'_1, \alpha'_2\} + m_i,$$

for each $(\alpha', s, m) \in D'$, each $i \in \{1,2\}$ and $t_i \in \{L, H\}$. The status-quo is $d^* = 0$. Consider a mechanism (α', s, m') that is incentive efficient. Feasibility implies that

$$\sum_{i=1}^{2} \sum_{t_i \in \{L,H\}} U_i'(\alpha', s, m'|t_i) = \sum_{t \in \{L,H\}^2} \frac{1}{2} \sum_{i=1}^{2} u_i'((\alpha'(t), s(t), m'(t)), t) \leq 60$$

Since utilities are independent of the types, incentive constraints imply that the two types of each agent expects identical utility gains. Hence, if a mechanism (α', s, m') is incentive efficient, then there exists $(x_1', x_2') \in \mathbb{R}^2$ such that $x_1' + x_2' = 0$, and

$$U_i'((\alpha', m')|L) = U_i'((\alpha', m')|H) \leq 15 + x_i',$$

for $i = 1, 2$. The inequality must in fact be binding, since any such interim payoffs can be achieved by an incentive compatible mechanism (α', s, m'), where $\alpha_i'(t) = 10$, $s_i(t) = 1/2$, $m_i'(t) = x_i'$, for each $t \in \{L, H\}^2$ and each $i \in \{1, 2\}$.

Finally, suppose that the impartial third party can choose to allocate the individuals' time between the three fields:

$$D'' = \{(\alpha_1, \alpha_2, \alpha_1', \alpha_2', s_1, s_2 m_1, m_2) \in [0, 10]^4 \times [0, 1]^2 \times \mathbb{R}^2 |$$
$$s_1 + s_2 = 1, \alpha_1 + \alpha_1' \leq 10, \alpha_2 + \alpha_2' \leq 10, \text{ and } m_1 + m_2 \leq 0\}$$

and the utility functions are given by the following expression:

$$u_i''((\alpha, \alpha', s, m), t) = \pi_i(t_i)\alpha_i + 3s_i \min\{\alpha_1', \alpha_2'\} + m_i,$$

for each $(\alpha, \alpha', s, m) \in D''$, each $i \in \{1, 2\}$ and $t_i \in \{L, H\}$, with the convention $\pi_i(L) = 1$ and $\pi_i(H) = 2$, for each $i \in \{1, 2\}$. Notice that

$$\sum_{i=1}^{2} \sum_{t_i \in \{L,H\}} U_i''((\alpha, \alpha', m)|t_i) = \sum_{t \in \{L,H\}^2} \frac{1}{2} \sum_{i=1}^{2} u_i''((\alpha(t), \alpha'(t), s(t), m(t)), t) \leq 65,$$

for each $(\alpha, \alpha', s, m) \in \mathcal{F}(\mathcal{S}'')$, the last equality following from the fact that the maximal total surplus is 40 when both individuals' type is H and is 30 otherwise. Hence there is no way to find a feasible mechanism that gives an interim utility of at least $15 + x_1'$ and $15 + x_2'$ to the low-type individuals, and $20 + x_1$ and $20 + x_2$ to the high-type individuals, which contradicts MON, since $D \subseteq D''$, $D' \subseteq D''$, and both $u_i''(d, t) = u_i(d, t)$, for each $i \in I$, each $d \in D$, and each $t \in \{L, H\}^2$, and $u_i''(d', t) = u_i'(d', t)$, for each $i \in I$, each $d' \in D'$, and each $t \in \{L, H\}^2$. ■

I-EFF and MON would not be incompatible in quasi-linear environments under incomplete information in the absence of incentive constraints (e.g. as in the case of verifiable information as in Wilson (1978), de Clippel and Minelli (2004) and de Clippel (2007)). To illustrate this point,[14] suppose for instance that the output of both individuals could

[14] The general argument is left to the reader.

be observed at no cost in the first problem described in the proof of Theorem 1, and that contracts could be made contingent on that observed output. Then the social planner or the arbitrator could implement a mechanism that requires both individuals to work for ten hours on their own field, and for the individual with a high output to pay \$10 to the individual with low output (if such a configuration occurs). The resulting interim utilities are 15 for both types of both individuals, which is also achievable in the last two problems described in the proof of Theorem 1, and hence I-EFF and MON are compatible in that example. Though the presence of incentive constraints is an important factor for the incompatibility of I-EFF and MON, it is important to emphasize the other factor, which is the presence of incomplete information, i.e. asymmetric information at the time of making the collective choice. So the tension between I-EFF and MON would disappear if the collective decision to be implemented at the interim stage was made at the ex-ante stage (i.e. before the agents learn their private information). To illustrate this point,[15] notice indeed that there exists an ex-ante incentive efficient mechanism in each of the three problems discussed in the proof of Theorem 1 that lead to an expected payoff of \$15 to both players, and hence efficiency and monotonicity are compatible, even when incentive constraints are imposed, when considered at the ex-ante stage. Similarly, the impossibility result would not hold in the presence of moral hazard instead of adverse selection. Theorem 1 thus shares some similitude with Myerson and Satterthwaite (1983) impossibility result, in that the incompatibility of their two properties, namely ex-post efficiency and interim individual rationality, also requires the presence of both incentive constraints and incomplete information.[16]

5. PARTIAL AXIOMATIC RESULTS

Readers will find in the Introduction the motivation that justifies the weaker version of MON that I am about to define. Interim utilities are *transferable*[17] at an incentive efficient mechanism μ if, for all $(i,j) \in I \times I$ with $i \neq j$ and all $(t_i,t_j) \in T_i \times T_j$ that comes with positive probability ($p(t_i,t_j) > 0$), there exists a feasible mechanism ν such that

$$U_i(\nu|t_i) > U_i(\mu|t_i),\ U_j(\nu|t_j) < U_j(\mu|t_j), \text{ and}$$
$$(\forall k \in I)(\forall t_k \in T_k) : (k,t_k) \notin \{(i,t_i),(j,t_j)\} \Rightarrow U_k(\nu|t_k) \geq U_k(\mu|t_k).$$

[15]Again, the general argument is left to the reader.

[16]Indeed, if any mechanism could be implemented in Myerson and Satterthwaite's example without having to bother with incentive constraints, then any mechanism in Wilson's (1978) coarse core (which, he proved, is non-empty in even much more general exchange economies), is both interim individually rational and ex-post efficient (because of quasi-linearity). On the other hand, the first-best can be achieved in Myerson and Satterthwaite's example (and many other quasi-linear environments), even while imposing the incentive constraints, if decisions are made at the ex-ante stage (see e.g. d'Aspremont ans Gérard-Varet (1979, 1982)). It is then always possible to design some type-independent monetary transfer to meet the (ex-ante) individual rationality constraints.

[17]Cf. footnote 5.

Restricted Monotonicity (R-MON) *Let $\mathcal{S}$ and $\mathcal{S}'$ be two social choice problem. Suppose that $\mathcal{S}'$ differs from $\mathcal{S}$ only in that more collective decisions are available: $I = I'$, $D \subseteq D'$, $T_i = T_i'$, and $u_i(d,t) = u_i'(d,t)$, for each $i \in I$, each $d \in D$, and each $t \in T$. Let $\mu \in \Sigma(\mathcal{S})$ be such that interim utilities are transferable at μ, and let $\mu' \in \Sigma(\mathcal{S}')$. Then $U_i(\mu'|t_i) \geq U_i(\mu|t_i)$, for each $t_i \in T_i$, and each $i \in I$.*

R-MON remains silent when it comes to mechanisms $\mu \in \Sigma(\mathcal{S})$ at which interim utilities are not transferable, contrary to MON. It is thus weaker than MON.

R-MON is now compatible with I-EFF. I will now add three axioms in the spirit of Kalai (1977, Theorem 1) in order to obtain a partial characterizarion of a natural extension of the egalitarian principle. The first property requires the social choice function to be covariant with respect to renaming the individuals and/or their types, as well as relabeling collective decisions.

Anonymity (AN) *Let $\mathcal{S} = (I, D, d^*, (T_i)_{i\in I}, p, (u_i)_{i\in I})$ and $\mathcal{S}' = (I, D, d^{**}, (T_i')_{i\in I}, p'$, $(u_i')_{i\in I})$ be two social choice problems. Suppose that there exist an isomorphism $f : I \to I$, an isomorphism $g : D \to D$ with $g(d^*) = d^{**}$, and isomorphisms $h_i : T_i \to T'_{f(i)}$ (one for each $i \in I$) such that*

1. *$(\forall t \in T) : p(t) = p'(h(t)))$, and*
2. *$(\forall t \in T)(\forall i \in I)(\forall d \in D) : u_i(d,t) = u'_{f(i)}(g(d), h(t))$,*

with the convention $h(t) = (h_i(t_i))_{i\in I}$. Then $\mu' \in \Sigma(\mathcal{S}')$ if and only if $\mu \in \Sigma(\mathcal{S})$, where μ is the mechanism for $\mathcal{S}$ defined as follows: the probability of implementing $d \in D$ when first individual reports $t_1 \in T_1$, ..., and the I^{th} individual reports $t_I \in T_I$ is equal to the probability of implementing $g(d)$ under μ' when individual $f(1)$ reports $h_1(t_1)$, ..., and individual $f(I)$ reports $h_I(t_I)$.

Most results in social choice under complete information, including Kalai's, are phrased in the space of utilities. This is sometimes referred to as the welfarist assumption. Following the same approach is not exactly feasible under incomplete information, as one needs to know the underlying decisions in order to have the possibility of phrasing the incentive constraints. Yet, one can impose an axiom in that spirit after taking these constraints into account. Readers will find a more detailed motivation for the next two axioms in the Introduction.

Interim Welfarism (I-WELF) *Let $\mathcal{S} = (I, D, d^*, (T_i)_{i\in I}, p, (u_i)_{i\in I})$ and $\mathcal{S}' = (I', D', d'$, $(T_i')_{i\in I}, p', (u_i')_{i\in I})$ be two social choice problems. If $T_i = T_i'$, for each $i \in I$, and $\mathcal{U}(\mathcal{F}(\mathcal{S})) = \mathcal{U}(\mathcal{F}(\mathcal{S}'))$, then $\mathcal{U}(\Sigma(\mathcal{S})) = \mathcal{U}(\Sigma(\mathcal{S}'))$.*

Of course, this definition boils down to the usual notion of welfarism under complete information, i.e. when each type set is a singleton.

The other axiom associated to the welfarist assumption requires that if a feasible mechanism generates the same interim utilities as another mechanism in the solution of a problem, then it also belongs to the solution of that problem.

Exhaustivity (EX) *Let $\mathcal{S} = (I, D, d^*, (T_i)_{i\in I}, p, (u_i)_{i\in I})$ be a social choice problem. If $\mu \in \Sigma(\mathcal{S})$ and μ' is a feasible mechanism such that $U_i(\mu'|t_i) = U_i(\mu|t_i)$, for all $t_i \in T_i$ and all $i \in I$, then $\mu' \in \Sigma(\mathcal{S})$.*

I now define the interim lex-min solution that is partially characterized below, and that boils down to the usual egalitarian criterion under complete information. A mechanism $\mu \in \mathcal{F}(\mathcal{S})$ belongs to the *interim lex-min solution* of the social choice problem $\mathcal{S} = (I, D, d^*, (T_i)_{i\in I}, p, (u_i)_{i\in I})$, $\mu \in \Sigma^{lex}(\mathcal{S})$, if and only if $\theta(\mathfrak{u}(\mu))$ maximizes $\theta(\mathfrak{u})$ according to the lexicographic order over all $\mathfrak{u} \in \mathcal{U}(S)$, where $\theta : \times_{i\in I}\mathbb{R}_+^{T_i} \to \times_{i\in I}\mathbb{R}_+^{T_i}$ is the function that rearrange the components of a vector increasingly.

Theorem 2 *The interim lex-min solution Σ^{lex} satisfies I-EFF, AN, R-MON, I-WELF, and EX. In addition, if Σ is a solution that satisfies the five axioms, and $\mathcal{S} = (I, D, d^*, (T_i)_{i\in I}, p, (u_i)_{i\in I})$ is a social choice problem such that*

1. *$\#T_i = \#T_j$, for all i, j, and*

2. *there exists $\mu \in \Sigma^{lex}(\mathcal{S})$ at which interim utilities are transferable,*

then $\Sigma(\mathcal{S}) = \Sigma^{lex}(\mathcal{S})$.

Proof: Let $\mathcal{S} = (I, D, d^*, (T_i)_{i\in I}, p, (u_i)_{i\in I})$ be a social choice problem. Observe that, if interim utilities are transferable at $\mu \in \Sigma^{lex}(\mathcal{S})$, then $U_i(\mu|t_i) = U_j(\mu|t_j)$, for all $t_i \in T_i$, all $t_j \in T_j$, and all $i, j \in I$. Indeed, otherwise, there exist (i, t_i) and (j, t_j) with $i \neq j$ such that $U_i(\mu|t_i) < U_j(\mu|t_j)$. Since interim utilities are transferable at μ, there exists a mechanism $\mu' \in \mathcal{F}(S)$ such that $U_i(\mu'|t_i) > U_i(\mu|t_i)$, $U_j(\mu'|t_j) < U_j(\mu|t_j)$, and $U_k(\mu'|t_k) \geq U_k(\mu|t_k)$, for all (k, t_k) different from (i, t_i) and (j, t_j). For each $\epsilon \in]0, 1[$, $\mu^\epsilon := \epsilon\mu' + (1-\epsilon)\mu \in \mathcal{F}(\mathcal{S})$. For ϵ small enough, $\theta(\mathfrak{u}(\mu^\epsilon))$ strictly dominates $\theta(\mathfrak{u}(\mu))$ according to the lexicographic order, thereby contradicting $\mu \in \Sigma^{lex}(\mathcal{S})$. So it must be indeed that all the components of $\mathfrak{u}(\mu)$ are identical. Consider now a larger social choice problem $\mathcal{S}'$, as defined in R-MON. Since $\mu \in \mathcal{F}(\mathcal{S}')$, it must be that the smallest component of $\mathfrak{u}(\mu')$ is larger of equal to the identical components of $\mathfrak{u}(\mu)$, for each $\mu' \in \Sigma^{lex}(\mathcal{S}')$. Hence Σ^{lex} satisfies R-MON, as desired. The very definition of Σ^{lex} makes it straightforward to check that it satisfies the four other axioms.

Consider now a solution Σ that satisfies the five axioms, and let $\mathcal{S}$ be a social choice problem that satisfies the two assumption in the statement of the theorem. Let $\mu \in \Sigma^{lex}(\mathcal{S})$ such that interim utilities are transferable at μ, as in *2.* All the components of $\mathfrak{u}(\mu)$ are thus identical – cf. the previous paragraph. Let ξ be this common component. The reasoning to prove that $\Sigma(\mathcal{S}) = \Sigma^{lex}(\mathcal{S})$ requires the consideration of other related social choice problems. I start by defining some of them. The first alternative problem is obtained by changing both the common prior and the utility functions so as to keep the set of interim utilities that are achievable via feasible mechanisms unchanged: $S^1 = (I, D, d^*, (T_i)_{i\in I}, p^1, (u_i^1)_{i\in I})$, with p^1 being the uniform probability distribution on T, and $u_i^1(d, t) := \#(T_{-i})p(t_{-i}|t_i)u_i(d, t)$, for all (d, t) and all i. Notice that this modification does not change the individuals' interim evaluations of any mechanism since the products of the conditional probabilities with the state-contingent utilities remain constant (see Myerson, 1984a, Section 3). Hence $\mathcal{U}(\mathcal{F}(\mathcal{S}^1)) = \mathcal{U}(\mathcal{F}(\mathcal{S}))$, and $\mu \in \Sigma^{lex}(\mathcal{S}^1)$.

Since μ is incentive efficient, a simple separation argument implies that there exists $\lambda \in \times_{i\in I}\mathbb{R}_+^{T_i}$ such that μ maximizes $\sum_{i\in I}\sum_{t_i\in T_i}\lambda_i(t_i)U_i(\nu|t_i)$, over all $\nu \in \mathcal{F}(\mathcal{S}^1)$. Given that interim utilities are transferable at μ, all the components of λ must actually be strictly positive. Let W^λ be the value of that weighted sum when evaluated at μ. Following Myerson's virtual utility construction (see also Lemma 4 in the Appendix), it is possible to construct an auxiliary problem $\mathcal{S}^2 = (I, D^2, (T_i)_{i\in I}, p^1, (u_i^2)_{i\in I})$, where $D^2 = D \cup \{d_{i,t_i} | t_i \in T_i, i \in I\}$, $u_i^2(\cdot, t) = u_i^1(\cdot, t)$ on D, for all t and all i, and such that $\mathcal{U}(\mathcal{F}(\mathcal{S}^2))$ is the convex hull of the vectors 0 and $\mathfrak{u}^{i,t_i}$, for each $t_i \in T_i$ and each $i \in I$, where $\mathfrak{u}_j^{i,t_i}(t_j) = 0$, for all $(j, t_j) \neq (i, t_i)$, and $\mathfrak{u}_i^{i,t_i}(t_i) = W^\lambda/\lambda_i(t_i)$. Notice that $\mu \in \Sigma^{lex}(\mathcal{S}^2)$, since μ remains incentive efficient in that larger problem.

Lemma 5 in the Appendix shows that it is possible to define utility functions u^3 and, for each combination (i, t_i), a collective decision $\hat{d}_{i,t_i}$ such that

1. $U_i^3(\hat{d}_{i,t_i} | t_i) = W^\lambda/\lambda_i(t_i)$;

2. $U_i^3(\hat{d}_{i,t_i} | t_i') = 0$ if $t_i' \neq t_i$;

3. $U_j^3(\hat{d}_{i,t_i} | t_j) = 0$, for all $j \in N \setminus \{i\}$ and all $t_j \in T_j$;

4. $\sum_{j\in I} \frac{\lambda_j(t_j')}{p^1(t_j')} u_j^3(\hat{d}_{i,t_i}, t') \leq \xi \sum_{j\in I} \frac{\lambda_j(t_j')}{p^1(t_j')}$, for all $t' \in T$

Let $D^3 = \{d^*\} \cup \{\hat{d}_{i,t_i} | i \in I, t_i \in T_i\}$, and $S^3 = (I, D^3, d^*, (T_i)_{i\in I}, p^1, (u_i^3)_{i\in I})$. Clearly, $\mathcal{U}(\mathcal{F}(\mathcal{S}^3)) = \mathcal{U}((\mathcal{S}^2))$.

Let

$$0 < \epsilon < \min_{(i,t_i)} \frac{\lambda_i(t_i)}{\sum_{j\in I} \sum_{t_j \in T_j} \frac{\lambda_j(t_j)}{p^1(t_j)}},$$

and let

$$\mathcal{C} = \{x \in \mathbb{R}_+^I | (\forall i \in I) : x_i + \epsilon \sum_{j \in I\setminus\{i\}} x_j \leq \xi(1 + \epsilon(\#I - 1))\}.$$

Notice that $\mathcal{C}$ is included in the half-space $\{x \in \mathbb{R}^I | \sum_{i\in I} \frac{\lambda_i(t_i)}{p^1(t_i)} x_i \leq \xi \sum_{i\in I} \frac{\lambda_i(t_i)}{p^1(t_i)}\}$, for all $t \in T$. Indeed, given t, let $x \in \mathcal{C}$ and let $J = \{i \in I | x_i > \xi\}$. If $J = \emptyset$, then $\sum_{i\in I} \frac{\lambda_i(t_i)}{p^1(t_i)} x_i$ is clearly lower or equal to $\xi \sum_{i\in I} \frac{\lambda_i(t_i)}{p^1(t_i)}$. Suppose then that J is nonempty, and let i be an element of J. The i-inequality in the definition of $\mathcal{C}$ implies that $x_i - \xi \leq \epsilon \sum_{j\in I\setminus J}(\xi - x_j)$. Multiplying by $\lambda_i(t_i)$, and taking the sum over $i \in J$, one gets: $\sum_{i\in J} \frac{\lambda(t_i)}{p^1(t_i)}(x_i - \xi) \leq \sum_{j\in I\setminus J} \epsilon(\sum_{i\in J} \frac{\lambda_i(t_i)}{p^1(t_i)})((\xi - x_j) \leq \sum_{j\in I\setminus J} \epsilon(\sum_{i\in I} \frac{\lambda_i(t_i)}{p^1(t_i)})(\xi - x_j) \leq \sum_{j\in I\setminus J} \frac{\lambda_j(t_j)}{p^1(t_j)}(\xi - x_j)$, and hence $\sum_{i\in I} \frac{\lambda_i(t_i)}{p^1(t_i)} x_i \leq \xi \sum_{i\in I} \frac{\lambda_i(t_i)}{p^1(t_i)}$, as desired.

Let $\mathcal{S}^4 = (I, D^4, d^*, (T_i)_{i\in I}, p^1, (u_i^4)_{i\in I})$ be the social choice problem with $D^4 = \mathcal{C}^T$ and $u_i^4(d^4, t) = d_i^4(t)$, for all $d^4 \in D^4$, all $i \in I$, and all $t \in T$. Observe that the problem $\mathcal{S}^4$ is symmetric, and hence AN and (1) imply that the interim utility of any mechanism in the solution of that problem must give equal interim utility to all the players and whatever their private information. I-EFF and EX imply that the constant mechanism that selects $(\xi, \ldots, \xi)$ in $\mathcal{C}$, for all $t \in T$, belongs to $\Sigma(\mathcal{S}^4)$. It is easy to check that interim utilities are transferable at that constant mechanism. The combination of R-MON and EX imply that it also belong to $\Sigma(\mathcal{S}^5)$, where $\mathcal{S}^5 = (I, D^5, d^*, (T_i)_{i\in I}, p^1, (u_i^5)_{i\in I})$ is the social choice problem with $D^5 = D^3 \cup D^4$, $u_i^5(d^5,t) = u_i^3(d^5,t)$ if $d^5 \in D^3$ and $u_i^5(d^5,t) = u_i^4(d^5,t)$ if $d^5 \in D^4$, for all $d^5 \in D^5$, all $i \in I$, and all $t \in T$. Indeed, the way ϵ was chosen guarantees that the constant mechanism is incentive efficient in $\mathcal{S}^5$ (cf. 4. and the previous paragraph). EX implies that any feasible mechanism that gives the same vector of interim utilities also belongs to $\Sigma(\mathcal{S}^5)$. Let's choose one that can be expressed via lotteries on D^3, which is possible since the constant mechanisms that pick one of the decisions in D^3 independently of the individuals' reports generate all the extreme points of $\mathcal{U}^{eff}(\mathcal{F}(\mathcal{S}^5))$. R-MON and EX imply that it must remain a solution to $\mathcal{S}^3$. I-WELF imply that any mechanism in the solution of $\mathcal{S}^2$ must have the same interim utilities, and μ must thus belong to $\Sigma(\mathcal{S}^2)$, by EX. R-MON and EX imply that $\mu \in \Sigma(\mathcal{S}^1)$. Hence $\Sigma^{lex}(\mathcal{S}^1) \subseteq \Sigma(\mathcal{S}^1)$, or $\Sigma^{lex}(\mathcal{S}^1) = \Sigma(\mathcal{S}^1)$, by (1). I-WELF and EX allow to conclude that $\Sigma^{lex}(\mathcal{S}) = \Sigma(\mathcal{S})$, as desired. ■

When there is complete information, i.e. when the type sets are singletons, Theorem 2 is similar to Kalai (1977, Theorem 1), or more precisely Moulin's (1988, Theorem 3.2) variant. There, cooperative opportunities are phrased directly in the space of utilities, making I-WELF and EX part of the model itself. Assumption 1 in the previous theorem is trivially satisfied under complete information. "Minimal transferability" guarantees that assumption 2 is also satisfied, and that the egalitarian solution satisfies the analogue of MON. In the absence of such assumptions, some authors have succeeded in characterizing the interim lex-min solution in that framework (see Imai, 1983, and Chun and Peters, 1989). Unfortunately, I have not been able yet to extend these results to social choice problems under incomplete information.

I now propose a variant of Theorem 2 that will prove useful in applications. It is important to have a partial characterization result that does not involve the cardinality restriction on the type sets (see condition *1* in Theorem 2). To that end, I impose a new axiom that is similar to "Splitting Types," as first defined by Harsanyi and Selten (1972) (see also Axiom 6 in Weidner (1992)). Let $\mathcal{S} = (I, D, d^*, (T_i)_{i\in I}, p, (u_i)_{i\in I})$ be a social choice problem, and let t_j be a type of an individual j. The social choice problem $\mathcal{S}' = (I, D, d^*, (T_i')_{i\in I}, p', (u_i')_{i\in I})$ obtained from S by *splitting* type t_j is defined as follows: $T_i' = T_i$, for all $i \in I \setminus \{j\}$, $T_j' = (T_j \setminus \{t_j\}) \cup \{t_{j1}, t_{j2}\}$, $p'(t') = p(t')$, for all $t' \in T'$ such that $t_j' \notin \{t_{j1}, t_{j2}\}$, $p'(t_{j1}, t_{-j}) = p'(t_{j2}, t_{-j}) = p(t)/2$, for all $t_{-j} \in T_{-j}$, $u_i'(d,t') = u_i(d,t')$, for all $i \in I$, all $d \in D$, and all $t' \in T'$ such that $t_j' \notin \{t_{j1}, t_{j2}\}$, $u_i'(d, t_{j1}, t_{-j}) = u_i'(d, t_{j2}, t_{-j}) = u_i(d,t)$, for all $i \in I$, all $d \in D$, and all $t_{-j} \in T_{-j}$. So individual j's types has been splitted into two sub-types that have the same information and the same preferences as when of type t_j in the original problem. As for j's other types, nothing has changed, and for the other individuals, they see j's two new subtypes half as likely as t_j in the original problem, and as having the exact same properties as t_j otherwise. Such a split is thus an irrelevant change in the way to model the situation at hand,[18] and it should be inconsequential on the solution of the problem. This is indeed

[18]If a mechanism $\mu : T \to \Delta(D)$ is feasible for $\mathcal{S}$, then $\mu' : T' \to \Delta(D)$ is feasible for $\mathcal{S}'$, where μ' is defined as follows: $\mu'(t') = \mu(t')$, for all $t' \in T'$ such that $t_j' \notin \{t_{j1}, t_{j2}\}$, and $\mu'(t_{j1}, t_{-j}) = \mu'(t_{j2}, t_{-j}) = \mu(t)$, for all $t_{-j} \in T_{-j}$. Conversely, if a mechanism $\mu' : T' \to \Delta(D)$ is feasible for $\mathcal{S}'$, then

the content of the following axiom.

Irrelevant Splitting of a Type (IST) *Let $\mathcal{S} = (I, D, d^*, (T_i)_{i\in I}, p, (u_i)_{i\in I})$ be a social choice problem, and let $\mathcal{S}'$ be the problem derived from $\mathcal{S}$ by splitting individual j's type t_j into two types t_{j1} and t_{j2}. If $\mu \in \Sigma(\mathcal{S})$, then $\mu' \in \Sigma(\mathcal{S}')$, where $\mu'(t') = \mu(t')$, for all $t' \in T'$ such that $t'_j \notin \{t_{j1}, t_{j2}\}$, and $\mu'(t_{j1}, t_{-j}) = \mu'(t_{j2}, t_{-j}) = \mu(t)$, for all $t_{-j} \in T_{-j}$.*

The second change with respect to Theorem 2 is the introduction of an axiom of independence of irrelevant alternatives. It will allow to weaken a bit condition *2* in Theorem 2.

Independence of Irrelevant Alternatives (IIA) *Let $\mathcal{S}$ and $\mathcal{S}'$ be two social choice problem. Suppose that $\mathcal{S}'$ differs from $\mathcal{S}$ only in that more collective decisions are available: $I = I'$, $D \subseteq D'$, $T_i = T'_i$, and $u_i(d,t) = u'_i(d,t)$, for each $i \in I$, each $d \in D$, and each $t \in T$. If $\mu \in \Sigma(\mathcal{S}') \cap \mathcal{F}(\mathcal{S})$, then $\mu \in \Sigma(\mathcal{S})$.*

MON implies IIA. R-MON implies IIA only when interim utilities are transferable in the smaller problem. IIA is thus only a small addition to the list of axioms. It will be interesting to note that Σ^{lex} satisfies IIA, even though it does not satisfy MON.[19]

Theorem 2' *The interim lex-min solution Σ^{lex} satisfies I-EFF, AN, R-MON, I-WELF, EX, IST, and IIA. In addition, if Σ is a solution that satisfies the seven axioms, and $\mathcal{S} = (I, D, d^*, (T_i)_{i\in I}, p, (u_i)_{i\in I})$ is a social choice problem for which there exists $\mu \in \Sigma^{lex}(\mathcal{S})$ such that*

1. *there exists $\lambda \in \times_{i\in I}\mathbb{R}^{T_i}_{++}$ such that $\mu \in \arg\max_{\nu\in\mathcal{F}(\mathcal{S})} \sum_{i\in I}\sum_{t_i\in T_i} \lambda_i(t_i)U_i(\nu|t_i)$,*[20]
2. *$U_i(\mu|t_i) = U_j(\mu|t_j)$, for each $i \neq j$ and each $t \in T$.*

then $\Sigma(\mathcal{S}) = \Sigma^{lex}(\mathcal{S})$.

Proof: We already argued that Σ^{lex} satisfies the first five axioms. It is straightforward to check that it satisfies IIA. IST follows from the fact that the set of interim utilities that can be achieved in the splitted problem is the set of interim utilities that can be achieved in the original problem, except that the utility associated to the (j, t_j) component now appears twice (once for (j, t_{j1}) and once for (j, t_{j2}), cf. footnote 18).

$U_j(\mu'|t_{j1}) = U_j(\mu'|t_{j2})$ and the mechanism $\mu : T \to \Delta(D)$ is feasible, where $\mu(t') = \mu'(t')$, for all $t' \in T$ such that $t'_j \neq t_j$, and either $\mu(t) = \mu'(t_{j1}, t_{-j})$, for all $t_{-j} \in T_{-j}$, or $\mu(t) = \mu'(t_{j2}, t_{-j})$, for all $t_{-j} \in T_{-j}$.

[19] IIA captures a property of rationality on the part of the uninformed third party making the collective decision. A violation of IIA would require a strong argument to justify such behavioral irrationality. MON or R-MON, on the other hand, go beyond by imposing some principle of distributive justice, thereby narrowing the type of moral preference that this social planner or arbitrator is maximizing. Here, alternative properties might be meaningful as well, and leading to characterization of other moral preferences, or alternative characterizations of the same moral preferences.

[20] The interim lex-min solution selects incentive efficient mechanisms, and thus there necessarily exists $\lambda \in \times_{i\in I}\mathbb{R}^{T_i}_{+} \setminus \{0\}$ such that $\mu \in \arg\max_{\nu\in\mathcal{F}(\mathcal{S})} \sum_{i\in I}\sum_{t_i\in T_i} \lambda_i(t_i)U_i(\nu|t_i)$. Condition 1 in Theorem 2' is a technical condition that requires the existence of a strictly positive λ. It is most often satisfied. It is always satisfied, for instance, if D is finite.

Consider now a solution Σ that satisfies the seven axioms, and let $\mathcal{S}$ be a social choice problem that satisfies the two assumption in the statement of the theorem. Let $\mu \in \Sigma^{lex}(\mathcal{S})$. All the components of $\mathfrak{u}(\mu)$ are identical, by *2*. Let ξ be this common component. As for Theorem 2, the reasoning to prove that $\Sigma(\mathcal{S}) = \Sigma^{lex}(\mathcal{S})$ requires the consideration of other related social choice problems. Though not very important, it is easier to apply the results from the Appendix if all the type profiles come with a strictly positive probability, and so I will consider a first modified social choice problem $\mathcal{S}^1$ where the probability distribution is uniform, but the utility functions are redefined so as to keep the interim utilities unchanged - see the proof of Theorem 2. Particularly, $\mu \in \Sigma^{lex}(\mathcal{S}^1)$.

The next modified problem is obtained by splitting the types in $\mathcal{S}^1$ so as to obtain a new problem where all the individuals have the same number of possible types. Let $\tau = \max_{i\in I} \#T_i$, let $h_i : T_i \to \{1,\ldots,\tau\}$ be an injective function, for each $i \in I$, let $\bar{t} \in T$, let $g_i : \{1,\ldots,\tau\} \to T_i$ be defined as follows:

$$g_i(t'_i) = \begin{cases} h_i^{-1}(t'_i) & \text{if } t'_i \in Im(h_i) \\ \bar{t}_i & \text{otherwise,} \end{cases}$$

for each $t'_i \in \{1,\ldots,\tau\}$ and each $i \in I$, and let

$$\mathcal{S}^2 = (I, D, d^*, (T_i^2)_{i\in I}, p^2, (u_i^2)_{i\in I}),$$

where $T_i^2 = \{1,\ldots,\tau\}$, for each $i \in I$, $p^2(t^2) = p^1(g(t^2))/\Pi_{\{i\in I|g_i(t_i^2)=\bar{t}_i\}}(\tau + 1 - \#T_i)$, and $u_i^2(d,t^2) = u_i^1(d, g(t^2))$, for each $i \in I$, where $g(t^2) = (g_i(t_i^2))_{i\in I}$. So $\mathcal{S}^2$ differs from $\mathcal{S}^1$ only in that, for each $i \in I$, type $\bar{t}_i$ has been splitted sufficiently many times so that individual i has τ possible types. The splitted version of μ (iteration of the definition in IST – see also footnote 18) belongs to $\Sigma^{lex}(\mathcal{S}^2)$, and the associated vector of interim utilities is constant, with ξ being the common component.

Condition 1 implies that there exists $\lambda \in \times_{i\in I}\mathbb{R}_{++}^{T_i^2}$ such that μ maximizes

$$\sum_{i\in I}\sum_{t_i\in T_i^2} \lambda_i(t_i)U_i(\nu|t_i),$$

over all $\nu \in \mathcal{F}(\mathcal{S}^2)$. Let W^λ be the value of that weighted sum when evaluated at μ. As when moving from $\mathcal{S}^1$ to $\mathcal{S}^2$ in the proof of Theorem 2, it is possible to construct an auxiliary problem $\mathcal{S}^3 = (I, D^3, (T_i^2)_{i\in I}, p^2, (u_i^3)_{i\in I})$, where $D^2 = D \cup \{d_{i,t_i}|t_i \in T_i, i \in I\}$, $u_i^3(\cdot,t) = u_i^2(\cdot,t)$ on D, for all t and all i, and such that $\mathcal{U}(\mathcal{F}(\mathcal{S}^3))$ is the convex hull of the vectors 0 and $\mathfrak{u}^{i,t_i}$, for each $t_i \in T_i$ and each $i \in I$, where $\mathfrak{u}_j^{i,t_i}(t_j) = 0$, for all $(j,t_j) \neq (i,t_i)$, and $\mathfrak{u}_i^{i,t_i}(t_i) = W^\lambda/\lambda_i(t_i)$. Notice that the splitted version of μ belongs to $\Sigma^{lex}(\mathcal{S}^3)$, since it remains incentive efficient in that larger problem. The assumption of Theorem 2 are satisfied for $\mathcal{S}^3$, and hence the splitted version of μ belongs to $\Sigma(\mathcal{S}^3)$. This mechanism also belongs to $\Sigma(\mathcal{S}^2)$, by IIA. Hence any mechanism in $\Sigma(\mathcal{S}^2)$ generates a constant vector of interim utilities, with ξ being the common component.

Let $\nu \in \Sigma(\mathcal{S})$. I-WELF and EX implies that $\nu \in \Sigma(\mathcal{S}^1)$. IST implies that the splitted version of ν belongs to $\Sigma(\mathcal{S}^2)$. Condition (1) implies that the associated vector of interim utilities is constant, with each component being equal to ξ. Hence $U_i(\nu|t_i) = \xi = U_i(\mu|t_i)$, for all $t_i \in T_i$ and all $i \in I$. EX implies that $\mu \in \Sigma(\mathcal{S})$. Hence $\Sigma^{lex}(\mathcal{S}) \subseteq \Sigma(\mathcal{S})$. EX and (1) implies that $\Sigma^{lex}(\mathcal{S}) = \Sigma(\mathcal{S})$, as desired. ■

6. EGALITARIANISM AND UTILITARIANISM RECONCILED

The egalitarian solution requires the possibility of measuring the individuals' utility gains in some common units. Also, when such measurements are possible, the egalitarian principle often comes in conflict with other normative criteria, the most prominent alternative being the utilitarian principle. Various authors (see e.g. Harsanyi (1963), Shapley (1969), Yaari (1981)) showed that these two observations can be dealt with simultaneously in the following sense. There is a unique way to rescale the participants' utilities so that there exists a feasible option that is optimal according to both the egalitarian and the utilitarian criteria (in the rescaled utilities). In addition, any such option is optimal according to Nash' product criterion (either in the original or in the rescaled utilities, since that criterion is invariant to linear transformations). Nehring (2004) has argued that maximizing the ex-ante sum of the individuals' utilities is the natural extension of the utilitarian criterion to problems with incomplete information. It is also the criterion that was applied, for instance, in Myerson and Satterthwaite (1983) to select a specific incentive compatible after proving the incompatibility if ex-post efficiency, interim individual rationality and incentive compatibility. The analysis of the previous section suggests that Σ^{lex} is a natural extension of the egalitarian solution to frameworks with incomplete information. One may thus wonder whether there is a way, in each social choice problem, to make both criteria compatible by rescaling the participants' interim utilities. The next theorem answers positively. Interestingly, it turns out that it is not necessary anymore to resort to a lexicographic refinement of the egalitarian criterion, contrarily to the previous section, as perfect equality of all the interim utilities can be achieved in the rescaled problem. In addition, the coincidence of the extended utilitarian and egalitarian criteria occurs at mechanisms that maximize Harsanyi and Selten's (1972) weighted Nash product, where the weight of each type of each participant equals its marginal probability. Weidner (1992) obtained a direct axiomatic characterization of that solution in a similar framework as the one studied here, but under the assumption of independent types.

Theorem 3 *Let $\mathcal{S} = (I, D, d^*, (T_i)_{i\in I}, p, (u_i)_{i\in I})$ be a social choice problem, and let $\mu^* \in \mathcal{F}(\mathcal{S})$. Then*

$$\mu^* \in \arg\max_{\mu\in\mathcal{F}(\mathcal{S})} \Pi_{i\in I}\Pi_{t_i\in T_i}[U_i(\mu|t_i)]^{p(t_i)} \tag{3}$$

if and only if μ^* satisfies the two following conditions for some $\lambda \in \times_{i\in I}\mathbb{R}^{T_i}_{++}$:

1. $\mu^* \in \arg\max_{\mu\in\mathcal{F}(\mathcal{S})} \sum_{i\in I}\sum_{t_i\in T_i} p(t_i)U_i^\lambda(\mu|t_i)$,

2. $(\forall t \in T)(\forall i \in I)(\forall j \in I) : U_i^\lambda(\mu^*|t_i) = U_j^\lambda(\mu^*|t_j)$,

where $U_i^\lambda(\mu|t_i) := \lambda_i(t_i)U_i(\mu|t_i)$*, for each* $t_i \in T_i$ *and each* $i \in I$*.*

Proof: Let $W^* = \Pi_{i\in I}\Pi_{t_i\in T_i}[U_i(\mu^*|t_i)]^{p(t_i)}$. Since it is assumed that there exists at least one element of $\mathcal{U}(\mathcal{F}(\mathcal{S}))$ with only strictly positive components, it must be that $W^* > 0$ and $\{(U_i(\mu^*|t_i))_{t_i\in T_i, i\in I}\} >> 0$ under (3). The sets $\mathcal{U}(\mathcal{F}(\mathcal{S}))$ and

$$\{\mathfrak{u} \in \times_{i\in I}\mathbb{R}_+^{T_i}|\Pi_{i\in I}\Pi_{t_i\in T_i}\mathfrak{u}_i(t_i)^{p(t_i)} \geq W^*\}$$

are both closed and convex. Under (3), their intersection is the singleton $\{(U_i(\mu^*|t_i))_{t_i\in T_i, i\in I}\}$. Hence the separating hyperplane theorem implies that (3) is equivalent to the existence of a vector $l \in \times_{i\in I}\mathbb{R}^{T_i}$ for wich the two following conditions hold:

1. $\mu^* \in \arg\max_{\mu\in\mathcal{F}(\mathcal{S})}\sum_{i\in I}\sum_{t_i\in T_i} l_i(t_i)U_i(\mu|t_i)$,
2. l is proportional to the gradient of the curve $\{\mathfrak{u} \in \times_{i\in I}\mathbb{R}_+^{T_i}|\Pi_{i\in I}\Pi_{t_i\in T_i}\mathfrak{u}_i(t_i)^{p(t_i)} \geq W^*\}$ at $(U_i(\mu^*|t_i))_{t_i\in T_i, i\in I}$.

The second condition itself is equivalent to the existence of a strictly positive number α such that

$$l_i(t_i) = \frac{\alpha p(t_i)}{U_i(\mu^*|t_i)},$$

for all $t_i \in T_i$ and all $i \in I$. The result thus follows, by taking $\lambda_i(t_i) = l_i(t_i)/p(t_i)$, for all $t_i \in T_i$ and all $i \in I$. ■

7. APPLICATIONS

In this section, I illustrate the previous results in some detail in three quasi-linear examples. The first one is about sharing the output created by individuals with private information regarding their marginal productivity. The main purpose of that first example is to show how Theorem 2 can be applied, and to note that Σ^{lex} coincides with the ex-post egalitarian solution in problems where incentive constraints do not play a critical role (information is non-exclusive in the sense of Postlewaite and Schmeidler (1986)) and there is no aggregate uncertainty regarding the total profit to be shared. The second example is about the production of a public good and how to share its cost, while the third example is about fair terms of trade. Beyond learning by observing how Σ^{lex} behaves, these two examples will illustrate how Theorem 2' is an important complement to Theorem 2, but also that there are cases where neither apply. Theorem 3 will prove useful in finding the incentive compatible mechanism that maximizes Harsanyi and Selten's (1972) weighted Nash product. The emphasis in the following examples is placed on simplicity rather than generality.

Example 1 *Consider a cooperative of* 100 *individuals who collectively own a technology. The total profit when they perform at an effort level* $e = (e_i)_{i=1}^{100}$ *is*

$$f(e) = \sum_{i=1}^{100} \pi_i e_i,$$

where π_i is agent i's marginal productivity, which is known to him alone. On the other hand, it is commonly known that 30 *of them have a relatively high marginal productivity ($\pi_i(H) = 2$), while the remaining individuals' marginal productivity is rather low ($\pi_i(L) = 1$). The cost of performing an effort e is $e^2/20$, for all the individuals (independently of their marginal productivity). Formally, the social choice problem is $\mathcal{S} = (I, D, d^*, (T_i)_{i\in I}, p, (u_i)_{i\in I})$, where $I = \{1, \ldots, 100\}$, $D = \{(e, \alpha) \in \mathbb{R}^I_+ \times \mathbb{R}^I_+\}$, $d^* = (0, 0)$, $T_i = \{L, H\}$, for each $i \in I$, $p(t) = k/(100^2)$, where k is the number of type profiles with exactly* 30 *L's and* 70 *H's, and*

$$u_i((e, \alpha), t) = \alpha_i \sum_{i\in I} \pi_i(t_i)e_i - \frac{e_i^2}{20},$$

for each $(e, \alpha) \in D$ and each $t \in T$, where α_i thus denotes the proportion of the output going to individual i, for each $i \in I$.

Consider the mechanism (e, α), where

$$(e_i(t), \alpha_i(t)) = \begin{cases} (0, 0) & \text{if } \#\{j \in I | t_j = H\} \neq 30 \\ (10, \frac{14.5}{1300}) & \text{if } \#\{j \in I | t_j = H\} = 30 \text{ and } t_i = L \\ (20, \frac{29.5}{1300}) & \text{if } \#\{j \in I | t_j = H\} = 30 \text{ and } t_i = H \end{cases}$$

for each $i \in I$. This mechanism is incentive efficient (and also ex-post efficient). Also, $U_i((e, \alpha)|t_i) = 9.5$, for all $t_i \in T_i$ and all $i \in I$. In addition, it is the only mechanism that achieves these interim utilities, and hence $\Sigma^{lex}(\mathcal{S}) = \{(e, \alpha)\}$. Interim utilities are transferable at (e, α). Thus Theorem 2 applies, and any solution Σ that satisfies EFF, R-MON, AN, I-WELF, and EX must coincide with Σ^{lex}. The mechanism (e, α) also maximizes Harsanyi and Selten's (1972) weighted Nash product over $\mathcal{F}(\mathcal{S})$. This follows from Theorem 3, since it is easy to check that (e, α) maximizes $\sum_{i\in I} \sum_{t_i\in T_i} p(t_i)U_i^\lambda(\mu|t_i)$ over $\mu \in \mathcal{F}(\mathcal{S})$ if $\lambda_i(L) = \lambda_i(H) = 1$, for all $i \in I$.

Example 2 *Consider the following problem, as in Myerson (1979). Two individuals have the option to cooperate by investing in a public project that costs* \$100*, which is known to bring a satisfaction of* \$90 *to the second individual, and either a satisfaction of* \$30*, with probability p, or of* \$90*, with probability $1-p$, to the first individual. Formally, $I = \{1, 2\}$, $D = \{(x, m) \in \{0, 1\} \times \mathbb{R}^2 | m_1 + m_2 \leq -100x\}$, where $x = 0$ means that the project is not carried out, $x = 1$ means that the project is carried out, $d^* = (0, (0, 0))$, $T_1 = \{30, 90\}$,[21] $u_1((x, m), 30) = 30x + m_1$, $u_1((x, m), 90) = 90x + m_1$, and $u_2((x, m), 30) = u_2((x, m), 90) = 90x + m_2$.*

$\mathcal{U}(\mathcal{F}(\mathcal{S}))$ thus coincides with the set of vectors

$$(30\bar{x}(30)+\bar{m}_1(30), 90\bar{x}(90)+\bar{m}_1(90); 90(p\bar{x}(30)+(1-p)\bar{x}(90))+p\bar{m}_2(30)+(1-p)\bar{m}_2(90)),$$

[21] T_2 is ignored to make notations lighter. This is inconsequential, since T_2 is a singleton, the second individual having no private information.

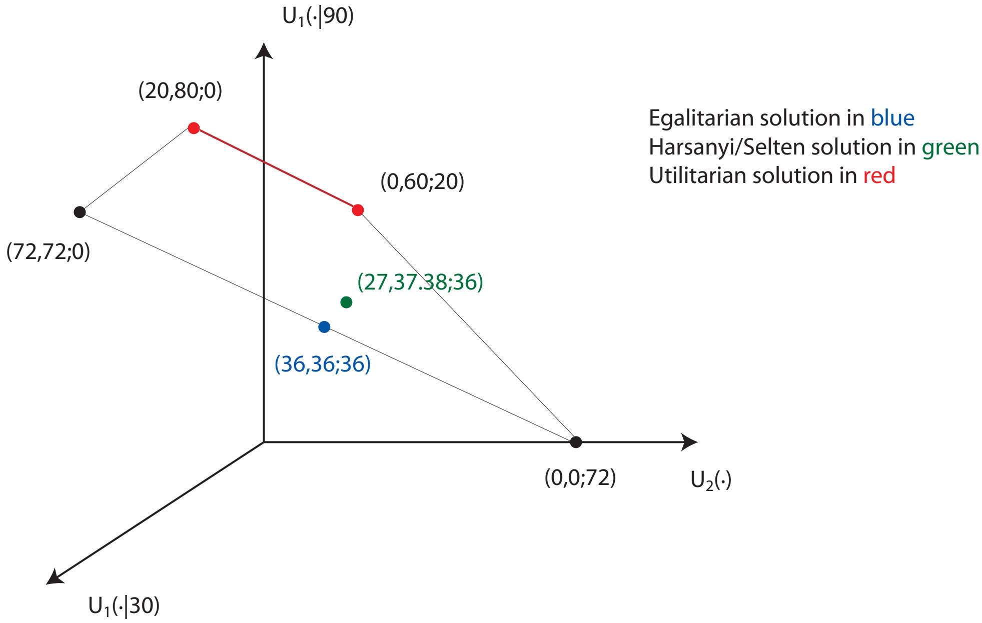

Figure 1: Illustration of Example 2 for p=1/10

that satisfy the following constraints:

$$\bar{m}_1(30) + \bar{m}_2(30) \leq -100\bar{x}(30) \text{ and } \bar{m}_1(90) + \bar{m}_2(90) \leq -100\bar{x}(90), \tag{4}$$

$$30(\bar{x}(90) - \bar{x}(30)) \leq \bar{m}_1(30) - \bar{m}_1(90) \leq 90(\bar{x}(90) - \bar{x}(30)), \tag{5}$$

where $\bar{x}(t_1)$ denotes the expected probability of implementing the public project when the first individual reports t_1, and $\bar{m}_i(t_1)$ is the expected monetary payoff received by individual i when the first individual reports t_1. Indeed, a feasible mechanism must select lotteries defined over D, and taking expectations of the feasibility constraints will give (4). Conversely, any vector $(\bar{x}(t_1), \bar{m}_1(t_1), \bar{m}_2(t_1)))$ can be seen as the expectation of the lottery that picks $[x = 0,\ m = (0,0)]$ with probability $1 - \bar{x}(t_1)$, and $[x = 1,$ $m = (\bar{m}_1(t_1)/\bar{x}(t_1), \bar{m}_2(t_1)/\bar{x}(t_1))]$ with probability $\bar{x}(t_1)$, which selects elements of D whenever (4) is satisfied. Inequalities in (5), on the other hand, represent the incentive contraints. A traditional reasoning allows to conclude that interim incentive efficiency implies that $\bar{x}(90) = 1$, and both inequalities in (4), as well as the second inequality of (5) are binding. If $p = 1/10$, as in Myerson (1979), then simple computations imply that $\mathcal{U}^{eff}(\mathcal{F}(\mathcal{S}))$ is the set of vectors

$$(30\bar{x}(30) + \bar{m}_1(30), 90\bar{x}(30) + \bar{m}_1(30); 72 - 82\bar{x}(30) - \bar{m}_1(30)), \tag{6}$$

with $\bar{x}(30) \in [0,1]$ and $\bar{m}_1(30) \in \mathbb{R}$. It is then easy to check that a feasible mechanism belongs to $\Sigma^{lex}(\mathcal{S})$ if and only if $\bar{x}(30) = 0$, $\bar{m}_1(30) = 36$, $\bar{m}_2(30) = -36$, $\bar{x}(90) = 1$, $\bar{m}_1(90) = -54$, and $\bar{m}_2(90) = -46$, in which case both individuals get an expected utility *of* 36*, independently of the true state. Theorem 2' implies that any solution that satisfies I-EFF, R-MON, AN, I-WELF, EX, IST, and IIA must coincide with* Σ^{lex} *in this numerical example. The ex-ante utilitarian solution, on the other hand, will contain only pooling mechanisms with* $\bar{x}(30) = \bar{x}(90) = 1$*,* $\bar{m}_1(90) = \bar{m}_1(30)$*, and* $\bar{m}_2(30) = \bar{m}_2(90) = -100 - \bar{m}_1(30)$*, leaving the choice of* $\bar{m}_1(30)$ *open. So, in this example, it does not refine the set of ex-ante incentive efficient mechanisms. The interim utilities are* $30+\bar{m}_1(30)$ *for the first individual when he values the public project at* 30*,* $90+\bar{m}_1(30)$ *for the first individual when he values the public project at* 90*, and* $90-\bar{m}_1(30)$ *in expectation for the second individual. One can apply Theorem 2 to find Harsanyi and Selten's (1972) weighted Nash product over* $\mathcal{U}(\mathcal{F}(\mathcal{S}))$*. One concludes from (6) that the vector* $(2/15, 13/15; 1)$ *is orthogonal to* $\mathcal{U}^{eff}(\mathcal{F}(\mathcal{S}))$*. If the weighted Nash optimum is reached at a point in the interior of* $\mathcal{U}^{eff}(\mathcal{F}(\mathcal{S}))$*, then it must be for* $\lambda = (20/15, 26/27, 1)$*, given the first condition in Theorem 2. Solving then for the two equations implied by the second condition in Theorem 2, one obtains* $\bar{x}(30) = 189/1092 \cong 0.173$*,* $\bar{m}_1(30) = 567/26 \cong 21.81$*,* $\bar{m}_2(30) = -1017/26 \cong -39.12$*,* $\bar{x}(90) = 1$*,* $\bar{m}_1(90) = -684/13 \cong -52.62$*, and* $\bar{m}_2(90) = -616/13 \cong -47.38$*, which turns out to be a feasible mechanism. The Harsanyi-Selten solution being unique, we are thus done solving the problem. The interim utilities are* 27 *for the first individual when he values the public project at* 30*,* $486/13 \cong 37.38$ *for the first individual when he values the public project at* 90*, and* 36 *in expectation for the second individual.*

The set of incentive efficient mechanisms and the three solutions are represented in the space of interim utilities on Figure 1. The utilitarian solution does not place any weight on the distribution of the gains from cooperation, and thereby allows to be more efficient in aggregate, the public project being implemented for sure independently of individual 1*'s type. On the other hand, this is achieved at the cost of equity, being too generous towards individual* 1 *when he values the public project at* \$90*, because one cannot rely on him to report his type for free. The second individual, for instance, cannot expect a gain larger than* \$20*, while the aggregate benefit is* \$80 *with probability* 9/10 *and* \$20 *with probability* 1/10*. The first individual's information rent vanishes at the egalitarian solution, in that both types of the first individual receive the same interim utility, but this comes at the cost of not implementing the public project when the first individual does not care much for it. Note that even though some mutually beneficial cooperative opportunities are not exploited, the mechanism itself, on the other hand, is incentive efficient. As already hinted by Theorem 2, the Harsanyi/Selten solution strikes a compromise between two conflicting points of view, allowing the public project to be realized with a positive probability when the first individual has the low type, but not systematically, so as to avoid being too soft on individual* 1 *when of the high type because of his informational advantage.*

Suppose now that there is a probability 9/10 *that the first individual values the public project at* \$30*, instead of* 1/10*. Starting from inequalities (4) and (5), it is easy to check that* $\mathcal{U}^{eff}(\mathcal{F}(\mathcal{S}))$ *is the set of vectors*

$$(30 + \bar{m}, 90 + \bar{m}; -10 - \bar{m}), \tag{7}$$

with $\bar{m} \in \mathbb{R}$*. In other words, incentive efficient mechanisms must be pooling, with the public project being implemented for sure regardless of the first individual's report, and* *the expected monetary compensation being constant. In this case, the first individual has a larger payoff when of the high type than when of the low type at any incentive efficient mechanism, and hence neither Theorem 2, nor Theorem 2', applies. The interim lex-min solution, on the other hand, seems to make sense. Since individual* 1 *systematically enjoys a payoff that is larger by a constant amount of* \$60 *when of a high type compared to the low type alternative, he is left aside and all what matters is to choose* m *so as to equalize the second individual's expected payoff with the first individual's payoff when he values the public project at* \$30*. Hence one must choose* $\bar{m} = -\$20$*, and the associated vector of interim utilities is* $(10, 70; 10)$*. The ex-ante utilitarian solution does not refine the set of incentive efficient mechanisms, while the Harsanyi-Selten solution selects a* $\bar{m}$ *that is slightly smaller than* $-\$20$*, thereby being slightly more generous towards the second individual.*[22]

More generally, Theorem 2' will apply whenever $p < 3/4$*, and the interim lex-min solution will prescribe to implement the public project if and only if the first individual reports his high type. In that case, he pays* \$90*, while the second individual pays* \$10*. In addition, a monetary transfer of* $\$(1-p)40$ *goes from the second to the first individual, independently of the type reported. The interim utility enjoyed by each type of each individual is* $(1-p)40$*. Theorems 1 and 1' will not apply when* $p \geq 3/4$*, in which case the interim lex-min solution has the public project implemented independently of the first individual's report, with* 1 *paying* 20% *of the cost, and* 2 *the remaining* 80%*.*

Example 3 *Consider the following bilateral trade problem, as in Myerson (1991, Section 10.3). A first individual owns one unit of a divisible good that is worth more to a second individual than to him. The good can be of relatively low quality, with probability* p*, in which case the good is worth* \$20 *per unit to the first individual and* \$30 *per unit to the second individual, or of relatively high quality, with probability* $1-p$*, in which case the good is worth* \$40 *per unit to the first individual and* \$50 *per unit to the second individual. The true quality of the object is known to the seller only. Cooperating here means agreeing on a quantity to trade against some monetary compensation, as a function of what the seller reveals about the quality of the good he owns. The problem is to find a fair compensation scheme. Formally,* $I = \{1, 2\}$*,* $D = \{(x, m) \in [0, 1] \times \mathbb{R}^2 | m_1 + m_2 \leq 0\}$*, where* x *represents the quantity traded,* $d^* = (0, (0, 0))$*,* $T_1 = \{L, H\}$*,*[23] $u_1((x, m), L) = m_1 - 20x$*,* $u_1((x, m), H) = m_1 - 40x$*,* $u_2((x, m), L) = 30x + m_2$*, and* $u_2((x, m), H) = 50x + m_2$*.*

$\mathcal{U}(\mathcal{F}(\mathcal{S}))$ *thus coincides with the set of vectors*

$$(m_1(L) - 20x(L), m_1(H) - 40x(H); p(30x(L) + m_2(L)) + (1-p)(50x(H) + m_2(H)))$$

[22]Not surprisingly, the Harsanyi-Selten solution converges towards the egalitarian solution, as p gets closer to 1, as the problem then approaches a problem with both complete information and transferable utilities, in which case the Nash and the egalitarian solutions coincide.

[23]As in the previous example, T_2 is ignored to make notations lighter. This is inconsequential, since T_2 is a singleton, the second individual having no private information.

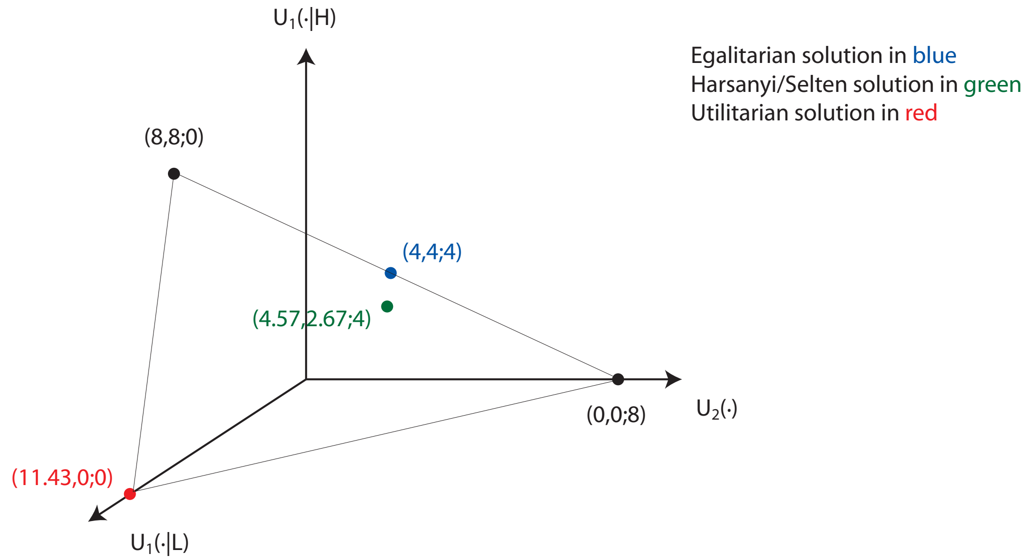

Figure 2: Illustration of Example 3 for p=4/5

that satisfy the following constraints:[24]

$$m_1(L) + m_2(L) \leq 0 \text{ and } m_1(H) + m_2(H) \leq 0, \tag{8}$$

$$20(x(L) - x(H)) \leq m_1(L) - m_1(H) \leq 40(x(L) - x(H)), \tag{9}$$

where $(x(t_1), m(t_1)) \in [0,1] \times \mathbb{R}$*, for each* $t_1 \in \{L, H\}$*. A traditional reasoning allows to conclude that interim incentive efficiency implies that* $x(L) = 1$*, and both inequalities in (8), as well as the first inequality of (9) are binding. If* $p = 4/5$*, as in Myerson (1991), then simple computations imply that* $\mathcal{U}^{eff}(\mathcal{F}(\mathcal{S}))$ *is the set of vectors*

$$(m - 20x, m - 40x; 8 + 26x - m), \tag{10}$$

with $(x, m) \in [0,1] \times \mathbb{R}$ *representing the quantity traded and the monetary compensation from* 2 *to* 1 *when the first individual reports* H*.* $\mathcal{U}^{eff}(\mathcal{F}(\mathcal{S}))$ *thus coincides with the triangle whose vertices are* $(80/7, 0; 0)$*,* $(8, 8, 0)$*, and* $(0, 0; 8)$*, as represented on Figure 2. The interim lex-min solution leads to the mechanism with* $(x, m) = (0, 4)$*, in which case both individuals enjoy an interim utility of* 4 *whatever their types (this remains true for any* $p > 1/3$*). Theorem 2' implies that any solution that satisfies I-EFF, R-MON, AN, I-WELF, EX, IST, and IIA must coincide with the interim lex-min solution in this numerical example. The ex-ante utilitarian solution, on the other hand, selects the*

[24]Utilities being linear in both the good and money, there is no loss of generality in discussing only deterministic mechanisms.

mechanism that is most advantageous to the first individual when he is of a low type - $(x, m) = (4/7, 160/7)$ leading to the extreme vector of interim utilities $(80/7, 0; 0)$. Contrarily to the previous example, the ex-ante utilitarian principle does refine the set of feasible mechanisms that are ex-ante efficient (which leads to interim utilities along the segment that joins $(0, 0; 8)$ to $(80/7, 0; 0)$). One can apply Theorem 2 to find Harsanyi and Selten's (1972) weighted Nash product over $\mathcal{U}(\mathcal{F}(\mathcal{S}))$. One concludes from (10) that the vector $(7/10, 3/10; 1)$ is orthogonal to $\mathcal{U}^{eff}(\mathcal{F}(\mathcal{S}))$. If the weighted Nash optimum is reached at a point in the interior of $\mathcal{U}^{eff}(\mathcal{F}(\mathcal{S}))$, then it must be for $\lambda = (7/8, 3/2; 1)$, given the first condition in Theorem 2. Solving then for the two equations implied by the second condition in Theorem 2, one obtains the feasible mechanism corresponding to the combination $(x, m) = (2/21, 136/21)$. The Harsanyi-Selten solution being unique, we are thus done solving the problem. The interim utilities are $32/7 \cong 4.57$ for the first individual when of a low type, $56/21 \cong 2.67$ for the first individual when of a high type, and 4 in expectation for the second individual.

More generally, Theorem 2' applies whenever $p > 1/3$, and the interim lex-min solution prescribes trade in full against $\$34$ when the first individual's report is L, and no trade, but still with a transfer of $\$4$ from 2 to 1 when the report is H. The interim utilities are $(4, 4; 4)$. If, on the other hand, the high type is rather likely to occur ($p \leq 1/3$) then incentive efficiency occurs only at pooling mechanisms where full trade occurs independently of the first individual's report. In that case, the low type always gets a larger utility than the high type, and Theorem 2' will not apply. The interim lex-min solution then follows the pragmatic principle of equalizing the two remaining payoffs. This leads to trade in full with a transfer of $\$45 - 10p$ from 2 to 1, independently of the first individual's report.

8. CONCLUDING COMMENTS

I explained in Section 4 why informational constraints may imply the incompatibility of EFF and MON, even on the class of quasi-linear social choice problems. Weakening MON into R-MON, I presented in Section 5 two partial axiomatic characterizations for Σ^{lex} that are applicable only to social choice problems $\mathcal{S}$ for which there exists an incentive efficient mechanism μ such that $U_i(\mu|t_i) = U_j(\mu|t_j)$, for all $i \neq j$ and all $t \in T$. I will thus restrict the discussion in this section to those social choice problems.

Mechanisms in Σ^{lex} are obviously equitable in those problems in the following sense: an uninformed third party (social planner or arbitrator) can be sure that all the individuals enjoy the same expected benefits given their own private information, and this whatever the actual profile of types. Example 2 was particularly simple to solve because information is non-exclusive (implying that first-best efficiency can be achieved via incentive compatible mechanisms), and there is no aggregate uncertainty regarding the number of low and high types in the population. In that case, Σ^{lex} coincides with the ex-post egalitarian solution. One might have another equity criterion in mind, for instance that people have the right to consume a quantity that is somehow linked to what they have produced. This is a classical discussion in Distributive Justice that, of course, remains present under incomplete information as well. The purpose of the present paper is strictly limited to the extension of the egalitarian principle to problems with asymmetric information. In that case, any member of the society has an equal right over the efficient surplus that can be created, independently of the role he or she has played in the creation of that surplus. Extending other principles, such as the Shapley value, for instance, is an interesting topic for further research (see Myerson (1984b) for a first definition, and the subsequent discussion in de Clippel's (2004)). A natural direction to go after this paper would be to extend Kalai and Samet (1985) to problems with asymmetric information.

I hope that the present paper will serve as a benchmark for future work in social choice under incomplete information. I see Σ^{lex} as the most straightforward extension of the egalitarian principle. While it does capture obvious equity considerations, as discussed in the previous paragraph, I believe that there might be alternative appealing ways to proceed. Here is an example of solution that takes more into account what the individuals know to determine what is equitable. I will restrict attention to the class of quasi-linear problems. Given his own information, an individual i of type t_i can evaluate the total surplus achieved by an incentive compatible mechanism μ:

$$TS_{i,t_i}(\mu) = \sum_{t_{-i} \in T_{-i}} p(t_{-i}|t_i) \sum_{i \in I} u_i(\mu(t), t).$$

From his point of view, his share of the total surplus realized by μ is then

$$s_{i,t_i}(\mu) = \frac{U_i(\mu|t_i)}{TS_{i,t_i}(\mu)}$$

(with the convention $s_{i,t_i}(\mu) = 1/\#I$ if both $U_i(\mu|t_i)$ and $TS_{i,t_i}(\mu)$ are equal to zero). One may have to accept inefficiency in some type profiles in order to satisfy the incentive constraints. For instance, the public good is not always implemented at incentive efficient mechanisms in Example 3, when $t_1 = 30$ and $p < 3/4$, and trade does not always occur at incentive efficient mechanisms in Example 4, when $t_1 = 40$ and $p > 1/3$. More generally, we know from Myerson and Satterthwaite (1983) that ex-post efficiency, interim individual rationality and incentive compatibility may be incompatible. This is why it is more natural for the definition of the total surplus to be endogenous to the mechanism considered, instead of taking the maximal total surplus that could be achieved in the absence of incentive constraints. Another meaningful definition of egalitarianism under incomplete information would then be to try find an interim individually rational and incentive efficient mechanism μ that equalizes the shares in each type profile: $s_{i,t_i}(\mu) = s_{j,t_j}(\mu)$, for all $i, j \in I$ and all $t \in T$. Perfect equalization being not always possible, as with Σ^{lex}, the general definition of this alternative criterion is to maximize, according to the lexicographic ordering, the vector $\theta(s(\mu))$ over the set of individually rational and incentive efficient mechanisms μ. The associated solution will be denoted Σ^*. Notice that, if an interim individually rational and incentive efficient mechanism μ for a social choice problem $\mathcal{S}$ is such that $s_{i,t_i}(\mu) = s_{j,t_j}(\mu)$, for all $i, j \in I$ and all $t \in T$, then $\mu \in \Sigma^*(\mathcal{S})$ and $s_{i,t_i}(\mu) = 1/\#I$, for all $i \in I$ and $t_i \in T_i$. Indeed, if σ is the share that is common to all the individuals of any type, then we have:

$$
\begin{aligned}
\textstyle\sum_{t\in T} p(t) \sum_{i\in I} u_i(\mu(t),t) &= \textstyle\sum_{i\in I}\sum_{t_i\in T_i} p(t_i) U_i(\mu|t_i) \\
&= \textstyle\sum_{i\in I}\sum_{t_i\in T_i} p(t_i)\sigma TS_{i,t_i}(\mu) \\
&= \sigma(\#I) \textstyle\sum_{t\in T} p(t)\sum_{i\in I} u_i(\mu(t),t),
\end{aligned}
$$

which implies $\sigma = 1/\#I$. Any mechanism μ' such that $\theta(s(\mu'))$ lexicographically dominates $\theta(s(\mu))$ will lead to similar equations, except that the second equality is now changed into a strict inequality for $\sigma = 1/\#I$, thereby leading to a contradiction and showing that $\mu \in \Sigma^*(\mathcal{S})$. It also implies, conversely, that any interim individually rational and incentive efficient mechanism μ such that $s_{i,t_i}(\mu) = 1/\#I$, for all $i \in I$ and $t_i \in T_i$, must belong to $\Sigma^*(\mathcal{S})$. Of course, Σ^* coincides with Σ^{lex} and the regular egalitarian criterion when information is complete, i.e. when the type sets contain a single element. The criterion also picks the same mechanism as Σ^{lex} in Example 2, because information is non-exclusive and there is no aggregate uncertainty. Easy computations in Example 3 show that, for any $p < 3/4$, equalization of the shares is feasible and a mechanism belongs to Σ^* if and only if $\bar{x}(30) = 4/7$, $\bar{m}(30) = (-80/7, -320/7)$, $\bar{x}(90) = 1$, and $\bar{m}(30) = (-50, -50)$. Things become even clearer if the collective decision implemented when $t_1 = 30$ is decomposed as an explicit lottery over D: the public project is implemented with probability 4/7, in which case player 1 pays \$20 and player 2 pays \$80 (this distribution of the cost is ex-post egalitarian), while there is no payment and no transfer when the public project is not implemented. The interim utilities are 40/7 for the first individual of type $t_1 = 30$, 40 for the first individual of type $t_1 = 90$, and $40 - \frac{240}{7p}$ for the second individual. So, on the one hand, one could say that the mechanism is not equitable from the point of view of an uniformed third party (social planner or arbitrator), in that different individuals enjoy different levels of satisfaction, but on the other hand the mechanism looks equitable from the point of view of the individuals themselves, given their private information. One must resort to the lexicographic ordering when $p \geq 3/4$, in which case the public project is implemented regardless of the first individual's report, who has to pay $\$\frac{130-90p}{5-3p}$, while the second individual pays $\$\frac{370-210p}{5-3p}$. In Example 4, for any $p > 1/3$, equalization of the shares is feasible and a mechanism belongs to Σ^* if and only if $x(L) = 1$, $m(L) = (25, -25)$, $x(H) = 1/5$, and $\bar{m}(H) = (9, -9)$ (or, equivalently, the second individual pays \$45 per unit to the first conditional on him reporting 40 - again this is the ex-post egalitarian outcome). The interim utilities are 5 for the first individual of type $t_1 = L$, 1 for the first individual of type $t_1 = H$, and $1 + 4p$ for the second individual, which again is different from the prescription made by Σ^{lex}. One must resort to the lexicographic ordering when $p \leq 1/3$, in which case the good is traded in full regardless of the first individual's report and he receives $\$45 - 10p$ from the second individual, which, this time, happens to coincide with the outcome of Σ^{lex}.

The fact that various extensions seem sensible is a feature that makes the subject of social choices under incomplete information richer and more interesting. Additional and sharper axiomatic characterizations are needed to capture the essence of what distinguishes these various normative criteria. Since they all coincide in the special case of complete information, the key normative question one must address is how to treat information in determining what is equitable, and more specifically how to treat different types of a same individual. The interim lex-min solution (as well as the ex-ante utilitarian principle, and Harsanyi and Selten's (1972) weighted Nash product) seems to treat different types of individuals as different individuals in the way it is computed. Analogous treatments have already appeared in very different contexts, cf. the notion of "type-agent" introduced by Harsanyi (1967-68) to define the notion of Bayesian Nash equilibrium, that also played a key role in defining a notion of core under asymmetric information (de Clippel (2007)). Yet it seems that other inter-type compromises might have some normative appeal as well. This is the right place to mention a third notion of equity that would also coincides with egalitarianism in quasi-linear social choice problems under complete information (with risk-neutral individuals), but leads to different recommendations under incomplete information. Procedural justice offers an interesting alternative to the consequentialist approach that discusses equity exclusively in terms of the outcomes selected in various problems. What matters now is to let the individuals themselves select the social outcome via a procedure (or a game form) that is fair, in a broad sense of giving them equal opportunities. "Random Dictatorship," choosing with equal probability one of the individuals to act as a dictator, provides an example of such procedures, at least when the resulting equilibrium outcome is Pareto efficient (as it is under complete information if the problem is quasi-linear and individuals are risk-neutral). This example of fair procedure is a bit extreme, since the outcome when a dictator has been selected is clearly unfair, but at least all the individuals are in an equal position ex-ante. Notice that "Random Dictatorship" is also a corner-stone of Myerson's (1984a) theory of bargaining. Though trivial, there is an interesting equivalence under complete information between the egalitarian principle and the equilibrium outcome of that procedure in quasi-linear problems with risk-neutral individuals. This general equivalence breaks down when information is incomplete. A key insight from Myerson's work (1983 and 1984a) (see also Maskin and Tirole (1992)) is that being a dictator defines an implicit inter-type compromise in some problems, and I will simply observe now that this implicit compromise is incompatible with both Σ^{lex} and Σ^* in both Examples 3 and 4.[25] Indeed, it is easy to check in Example 3 that the principal-agent game when the first individual is a dictator admits a unique weak sequential equilibrium outcome, with the public project being implemented, the first individual paying \$10 and the second paying \$90, independently of the reported type. The associated interim utilities are $(20, 80; 0)$. The unique weak sequential equilibrium outcome when the second individual is the dictator and $p < 3/4$ implies that the project is implemented only when the reported type is 90, in which case the first individual pays \$90 and the second pays \$10. The associated interim utilities are $(0, 0; 80(1 - p))$. When $p > 3/4$, the public project is always implemented, with the first individual paying \$30 and the second paying \$70. The associated interim utilities are $(0, 60; 20)$. Random Dictatorship thus leads to the interim utilities $(10, 40, 40(1 - p))$ if $p < 3/4$ and $(10, 70, 10)$ if $p > 3/4$. We see that Σ^{lex} coincides with the random dictatorship outcome if and only if $p > 3/4$. Σ^* is different from both Σ^{lex} and the random dictatorship outcome for any p. In Example 4, if $p > 1/3$, then the unique weak sequential equilibrium outcome when the first individual is the dictator

[25] Random dictatorship leads to a similar outcome as both Σ^{lex} and Σ^* in Example 2.

leads to the good being traded at the highest possible price - \$30 per unit if the reported type is L, and \$50 per unit if the reported type if H - but only a third is traded when the reported type is H, while the good is traded in full when the reported type is H. If the second individual is the dictator with $p > 1/3$, then the good is traded if and only if the reported type is L. in which case it is traded in full at the lowest possible price - \$20 per unit. Random Dictatorship thus leads to the interim utilities $(5, 5/3, 5p)$, which differ from both Σ^{lex} and Σ^*. If $p < 1/3$, then trade occurs in full independently of the first individual's report. The price is $\$50 - 20p$ if the first individual is the dictator, and \$40 if the second individual is the dictator. Random Dictatorship thus leads to the interim utilities $(25 - 10p, 5 - 10p, 5 - 10p)$, which coincide with the outcome of both Σ^{lex} and Σ^*.

It is a good place to emphasize the strength of I-WELF, as neither Σ^*, nor the Random Dictatorship solution, satisfy it. To see that, notice first that $\mathcal{U}(\mathcal{F}(\mathcal{S}))$ in Example 4 with $p = 4/5$ is the convex hull of $(0, 0; 0)$ and the three vectors shown on Figure 2, i.e. $(11.43, 0; 0)$, $(8, 8; 0)$, and $(0, 0; 8)$. This follows immediately from the previous characterization of the incentive efficiency frontier and the fact that $U_1(\mu|L) \geq U_1(\mu|H)$ at any incentive compatible mechanism (indeed, $m_1(L) - 20x(L) \geq m_1(H) - 20x(H) \geq m_1(H) - 40x(H)$, where the first inequality follows from the incentive constraint and the second follows from the fact that $x(H)$ is non-negative). Now, similar computations would show that the set of interim utilities that can be achieved by mechanisms that are individually rational and incentive compatible remains unchanged if p is $3/4$ instead of $4/5$, the buyer's value in the low type is \$92/3 instead of \$30, and the buyer's value in the low type is \$44 instead of \$50. Yet Σ^* will now leads to the interim utilities $(16/3, 16/33; 136/33)$ instead of $(5, 1; 21/5)$, and Random Dictatorship will now prescribe $(11/3, 8/9; 11/4)$ instead of $(5, 5/3; 4)$.[26]

I conclude this section with directions for future research (in addition to those already mentioned in the previous paragraphs). First, the most specific, would be to obtain full characterizations of the interim lex-min solution, either by completing Theorems 1 and 1', or by adapting other axiomatic characterizations of the egalitarian principle under complete information (see e.g. the axiom of Step-by-Step Decomposition in Kalai (1977), or of Monotonicity with respect to Changes in the Number of Agents in Thomson (1983a)). Second, it would be useful to get characterizations of the interim lex-min solution on more resrictive domains, e.g. the class of quasi-linear social choice problems, and also to derive I-WELF instead of imposing it (see e.g. de Clippel and Bejan (2009) under complete information). Third, one would like to get a definition and an axiomatic characterization of the related scale-covariant principle of relative egalitarianism (see e.g. Kalai and Smorodinsky (1975) for an axiomatic characterization of a similar solution in bargaining theory under complete information, Thomson (1983b), and Sprumont (2009)).

[26] Σ^* leads to the mechanism with $x(L) = 1$, $m(L) = (76/3, -76/3)$, $x(H) = 8/33$, and $m(H) = 42x(H)$. If the first individual is a dictator, then the resulting equilibrium outcome corresponds to the mechanism with $x(L) = 1$, $m(L) = (92/3, -92/3)$, $x(H) = 4/9$, and $m(H) = 44x(H)$. If the second individual is a dictator, then the resulting equilibrium outcome corresponds to the mechanism with $x(L) = 1$, $m(L) = (20, -20)$, $x(H) = 0$, and $m(H) = 0$.

The question here is which reference point to use to define the relative gains (see also de Clippel et al. (2010) for a related definition of egalitarian equivalence under incomplete information). Fourth, one may wish to check whether characterization results derived in the case of weak Bayesian implementation (requiring that the social planner or the arbitrator has the ability to focalise the individuals on a specific equilibrium of the mechanism he is implementing) also hold for other informational constraints (e.g. full Bayesian implementation, or strategy-proofness).

References

Aumann, R. J., 2003. Presidential Address. *Games and Economic Behavior* **45**, 2-14.

Chun, Y., and H. Peters, 1989. Lexicographic Monotone Path Solutions. *OR Spektrum* **11**, 43-47.

d'Aspremont, C., and L.-A. Gérard-Varet, 1979. Incentives and Incomplete Information. *Journal of Public Economics* **11**, 25-45.

d'Aspremont, C., and L.-A. Gérard-Varet, 1982. Bayesian Incentive Compatible Beliefs. *Journal of Mathematical Economics* **10**, 83-103.

de Clippel, G., 2005. Values for Cooperative Games with Incomplete Information: An Eloquent Example. *Games and Economic Behavior* **53**, 73-82.

de Clippel, G., 2007. The Type-Agent Core for Exchange Economies with Asymmetric Information. *Journal of Economic Theory* **135**, 144-158.

de Clippel, G., 2010. A Comment on "The Veil of Public Ignorance." Brown University, Working Paper 2010-3.

de Clippel, G., and C. Bejan, 2009. No Profitable Decompositions in Quasi-Linear Allocation Problems. Brown University, Working Paper 2009-6.

de Clippel, G., and E. Minelli, 2004. Two-Person Bargaining with Verifiable Information. *Journal of Mathematical Economics* **40**, 799-813.

de Clippel, G., D. Perez-Castrillo, and D. Wettstein, 2010. Egalitarian Equivalence under Asymmetric Information. Brown University, Working Paper 2010-5.

Feldman, A., and A. Kirman, 1974. Fairness and Envy. *The American Economic Review* **64**, 995-1005.

Foley, D., 1967. Resource Allocation and the Public Sector. *Yale Economic Essays* **7**.

Ginés, M., and F. Marhuenda, 2000. Welfarism in Economic Domains. *Journal of Economic Theory* **93**, 191-204.

Guo, M., and V. Conitzer, 2009. Worst-Case Optimal Redistribution of VCG Payments in Multi-Unit Auctions. *Games and Economic Behavior* **67**, 69-98.

Harsanyi, J. C., 1963. A Simplified Bargaining Model for the n-Person Cooperative Game. *International Economic Review* **4**, 194-220.

Harsanyi, J. C., 1967-68. Games with Incomplete Information Played by Bayesian Players. *Management Science* **14**, 159-182, 320-334, 486-502.

Harsanyi, J. C., and R. Selten, 1972. A Generalized Nash Solution for Two-Person Bargaining Games with Incomplete Information. *Management Science* **18**, 80-106.

Holmström, B., and R. B. Myerson, 1983. Efficient and Durable Decision Rules with Incomplete Information. *Econometrica* **51**, 1799-1819.

Imai, H., 1983. Individual Monotonicity and Lexicographic Maxmin Solution. *Econometrica* **51**, 389-402.

Kalai, E., 1977. Proportional Solutions to Bargaining Situations: Interpersonal Utility Comparisons. *Econometrica* **45**, 1623-1630.

Kalai, E., and D. Samet, 1985. Monotonic Solutions to General Cooperative Games. *Econometrica* **53**, 307-327.

Kalai, E., and M. Smorodinsky, 1975. Other Solutions to Nash's Bargaining Problem. *Econometrica* **43**, 513-518.

Luce, R.D., Raiffa, H., 1957. Games and Decisions. Wiley, New York.

Maskin, E. and J. Tirole, 1992. The Principal-Agent Relationship with an Informed Principal. II: Common Values. *Econometrica* **60**, 1-42.

Maniquet, F., and Y. Sprumont, 2010. Sharing the Cost of a Public Good: An Incentive-Constrained Axiomatic Approach. *Games and Economic Behavior* **68**, 275-302.

Mirlees, J. A., 1971. An Exploration in the Theory of Optimum Income Taxation. *Review of Economic Studies* **38**, 175-208.

Moulin, H., 1985. Egalitarianism and Utilitarianism in Quasi-Linear Bargaining. Econometrica **53**, 49-67.

Moulin, H., 1988. Axioms of Cooperative Decision Making. Cambridge University Press.

Moulin, H., 1994. Serial Cost-Sharing of Excludable Public Goods. *Review of Economic Studies* **61**, 305-325.

Moulin, H., 2009. Almost Budget-Balanced VCG Mechanisms to Assign Multiple Objects. *Journal of Economic Theory* **144**, 96-119.

Moulin, H., 1992. An application of the Shapley value to fair division with money. *Econometrica* **60**, 1331-1349.

Moulin, H., and W. Thomson, 1988. Can Everyone Benefit from Growth?: two Difficulties. *Journal of Mathematical Economics* **17**, 339-345.

Myerson, R. B., 1979. Incentive Compatibility and the Bargaining Problem. *Econometrica* **47**, 62-73.

Myerson, R. B., 1983. Mechanism Design by an Informed Principal. *Econometrica* **51**, 1767-1797.

Myerson, R. B., 1984a. Two-Person Bargaining Problems with Incomplete Information. *Econometrica* **52**, 461-487.

Myerson, R. B., 1984b. Cooperative Games with Incomplete Information. *International Journal of Game Theory* **13**, 69-96.

Myerson, R.B., 1991. Game Theory: Analysis of Conflict. Harvard University Press, Cambridge, MA.

Myerson, R. B. and M. Satterthwaite, 1983. Efficient Mechanisms for Bilateral Trading. *Journal of Economic Theory* **23**, 265-281.

Nehring, K., 2004. The Veil of Public Ignorance. *Journal of Economic Theory* **119**, 247-270.

Olszewski, W., 2004. Coalition Strategy-Proof Mechanisms for Provision of Excludable Public Goods. *Games and Economic Behavior* **46**, 88-114.

Pazner, E., and D. Schmeidler, 1978. Egalitarian Equivalent Allocations: a New Concept of Economic Equity. *Quarterly Journal of Economics* **92**, 671-687.

Postlewaite, A., and D. Schmeidler, 1986. Implementation in Differential Information Economies. *Journal of Economic Theory* **39**, 14-33.

Roemer, J. E., 1986. The Mismarriage of Bargaining Theory and Distributive Justice. *Ethics* **97**, 88-110.

Shapley, L. S., 1969. Utility Comparisons and the Theory of Games. In: Guilbaud, G. Th. (Ed.), La Decision, CNRS, 251-263.

Sprumont, Y., 2009. Relative Egalitarianism and Related Criteria. CIREQ Cahier 02-2009.

Thomson, W., 1983a. Problems of Fair Division and the Egalitarian Solution. *Journal of Economic Theory* **31**, 211-226.

Thomson, W., 1983b. The Fair Division of a Fixed Supply among a Growing Population. *Mathematics of Operation Research* **8**, 319-326.

Thomson, W., and R. B. Myerson. Monotonicity and Independence Axioms. *International Journal of Game Theory* **9**, 37-49.

Varian, H., 1974. Equity, Envy and Efficiency. *Journal of Economic Theory* **9**, 63-91.

Weidner, F., 1992. The Generalized Nash bargaining Solution and Incentive Compatible Mechanisms. *International Journal of Game Theory* **21**, 109-129.

Wilson, R., 1978. Information, Efficiency, and the Core of an Economy. *Econometrica* **46**, 807816.

Yaari, M. E., 1981. Rawls, Edgeworth, Shapley, Nash: Theories of Distributive Justice Re-Examined. *Journal of Economic Theory* **24**, 1-39.

Young, P. H., 1985a. Producer Incentives in Cost Allocation. *Econometrica* **53**, 757-765.

Young, P. H., 1985b. Monotonic Solutions of Cooperative Games. *International Journal of Game Theory* **14**, 65-72.

APPENDIX

Results in this section are variants or reformulations of previous results established by Myerson (1983, 1984a, 1984b, 1991). It is thus not meant for publication.

Lemma 1 *Let $(T_i)_{i\in I}$ be finite sets of types, let p be a probability distribution on $T = \times_{i\in I} T_i$ with full support, let $v : T \to \mathbb{R}$, and let $\mathfrak{u} \in \times_{i\in I}\mathbb{R}^{T_i}$. Then there exists $x : T \to \mathbb{R}$ such that*

1. $(\forall t \in T) : \sum_{i\in I} x_i(t) = v(t)$, *and*
2. $(\forall i \in I)(\forall t_i \in T_i) : \mathfrak{u}_i(t_i) = \sum_{t_{-i}} p(t_{-i}|t_i)x_i(t)$,

if and only if

$$\sum_{i\in I}\sum_{t_i\in T_i} p(t_i)\mathfrak{u}_i(t_i) = \sum_{t\in T} p(t)v(t).$$

<u>Proof</u>: If the two conditions are true, then

$$\begin{aligned}\textstyle\sum_{i\in I}\sum_{t_i\in T_i} p(t_i)\mathfrak{u}_i(t_i) &= \textstyle\sum_{i\in I}\sum_{t_i\in T_i} p(t_i)\sum_{t_{-i}} p(t_{-i}|t_i)x_i(t)\\ &= \textstyle\sum_{t\in T} p(t)\sum_{i\in I} x_i(t)\\ &= \textstyle\sum_{t\in T} p(t)v(t),\end{aligned}$$

where the first equality follows from *2*, the second equality follows from rearranging the terms, and the third equality follows *1*.

I now prove the converse, assuming for simplicity that $\sum_{t\in T} p(t)v(t) \neq 0$ (a straightforward translation argument implies that the result also holds when $\sum_{t\in T} p(t)v(t) = 0$). To do this, I will show that for any $(i, \bar{t}_i)$ there exists an x that satisfies *1* and *2* for $\mathfrak{u}^{i.\bar{t}_i}$, where $\mathfrak{u}_i^{i.\bar{t}_i}(t_i) = \sum_{t\in T} p(t)v(t)/p(\bar{t}_i)$ and $\mathfrak{u}_j^{i.\bar{t}_i}(t_j) = 0$, for all $(j, t_j) \neq (i, \bar{t}_i)$. The result will indeed follow since any vector $\mathfrak{u}$ such that $\sum_{i\in I}\sum_{t_i\in T_i} p(t_i)\mathfrak{u}_i(t_i) = \sum_{t\in T} p(t)v(t)$ can be written as an affine combination of these vectors, and the equations in *1* and *2* are linear in x. So I now define a function $x : T \to \mathbb{R}^I$, and show that it satisfies the two sets of equations for $\mathfrak{u}^{i.\bar{t}_i}$:

$$x_i(t) = 0 \text{ and } x_j(t) = \frac{v(t)}{\#I - 1},$$

for all $t \in T$ such that $t_i \neq \bar{t}_i$, and

$$x_i(\bar{t}_i, t_{-i}) = v(\bar{t}_i, t_{-i}) - \sum_{j\in N\setminus\{i\}} x_j(t)$$

$$x_j(\bar{t}_i, t_{-i}) = -\frac{1}{\#I - 1}\sum_{\tilde{t}_{-j}|\tilde{t}_i\neq\bar{t}_i} \frac{p(\tilde{t}_{-j}, t_j)}{p(\bar{t}_i, t_j)}v(t).$$

for all $t_{-i} \in T_{-i}$. Equations in *1* and the equations in *2* for any (i, t_i) with $t_i \neq \bar{t}_i$ are trivially satisfied, by construction. I now check the equations in *2* for any (j, t_j) with $j \neq i$. For each $j \in I \setminus \{i\}$ and each $t_j \in T_j$, we have:

$$\sum_{t_{-j}} p(t_{-j}|t_j)x_j(t) = \sum_{t_{-ij}} p(t_{-ij}, \bar{t}_i|t_j)x_j(\bar{t}_i, t_{-i}) + \sum_{t_{-j}|t_i \neq \bar{t}_i} p(t_{-j}|t_j)x_j(t)$$

$$= -\frac{p(\bar{t}_i|t_j)}{\#I - 1} \sum_{t_{-j}|t_i \neq \bar{t}_i} \frac{p(t_{-j}, t_j)}{p(\bar{t}_i, t_j)} v(t) + \frac{1}{\#I - 1} \sum_{t_{-j}|t_i \neq \bar{t}_i} p(t_{-j}|t_j)v(t),$$

where the second equality follows from the definition of $\bar{x}$, and in particular the fact that $\bar{x}_j(\bar{t}_i, t_{-i})$ does not depend on t_{-ij}. The last expression equals zero, as desired, because $p(\bar{t}_i|t_j)\frac{p(t_{-j},t_j)}{p(\bar{t}_i,t_j)} = p(t_{-j}|t_j)$. Finally, since x satisfies the equations in *1*, we have that $\sum_{i\in I}\sum_{t_i\in T_i} p(t_i)\sum_{t_{-i}} p(t_{-i}|t_i)x_i(t) = \sum_{t\in T} p(t)v(t)$. This combined with what we just proved implies that $\sum_{t_{-i}} p(t_{-i}|\bar{t}_i)x_i(t) = \sum_{t\in T} p(t)v(t)/p(\bar{t}_i)$, as desired. ■

Lemma 2 *A feasible mechanism μ is incentive efficient if and only if there exist $\lambda \in \times_{i\in I}\mathbb{R}_+^{T_i} \setminus \{0\}$ and $\alpha \in \times_{i\in I}\mathbb{R}_+^{T_i\times T_i}$ such that*

1. $(\forall i \in I)(\forall (t_i, t_i') \in T_i \times T_i) : \alpha_i(t_i'|t_i)(U_i(\mu|t_i) - U_i(\mu, t_i'|t_i)) = 0$

2. $\sum_{i\in I}\sum_{t_i\in T_i} \lambda_i(t_i)U_i(\mu|t_i) = \sum_{t\in T} p(t)\max_{d\in D}\sum_{i\in I} v_i^{(\lambda,\alpha)}(d, t)$*, where the "virtual utility" functions* $v_i^{(\lambda,\alpha)}$ *are defined as follows:*

$$v_i^{(\lambda,\alpha)}(d, t) = \frac{1}{p(t_i)}[(\lambda_i(t_i) + \sum_{t_i'} \alpha_i(t_i'|t_i))u_i(d, t) - \sum_{t_i'} \alpha_i(t_i|t_i')u_i(d, t_{-i}, t_i')],$$

for each $d \in D$, each $t \in T$, and each $i \in I$.

Proof: The vector λ is derived from a classical separation argument, using the fact that $\mathcal{U}(\mathcal{F}(\mathcal{S}))$ is closed and convex. The vector α specifies the dual variables of maximization of the weighted sum under the incentive constraints. Condition 1 is the usual condition stating that the dual variable associated to the constaint that individual i should not pretend to be of type t_i' when being actually of type t_i, is positive only if that constraint is binding. Condition 2 is obtained by rearranging the terms of the Lagrangean (more details available in Myerson (1991, Chapter 10), for instance). ■

Lemma 3 *Let $A = (a_{ij})_{1\leq i,j\leq n}$ be a square matrix with non-positive elements off the diagonal (i.e. $a_{ij} \leq 0$ if $i \neq j$) and such that the sum of the elements in each column is strictly positive (i.e. $\sum_{i=1}^n a_{ij} > 0$, for each j). Then A is invertible.*

Proof: Let $x \in \mathbb{R}^n$ be such that $Ax = 0$. I will prove that $x \geq 0$. Suppose, on the contrary, that $J = \{j \in I|x_j < 0\} \neq \emptyset$. Then

$$\sum_{j\in J}\sum_{k=1}^n a_{jk}x_k = 0,$$

which is equivalent to

$$\sum_{j\in J}\sum_{k\in J} a_{jk}x_k + \sum_{j\in J}\sum_{k\in I\setminus J} a_{jk}x_k = 0.$$

Notice that $\sum_{j\in J}\sum_{k\in J} a_{jk}x_k = \sum_{k\in J}(\sum_{j\in J} a_{jk})x_k < 0$, because $\sum_{j\in J} a_{jk} > 0$, for each $k \in J$, given the assumptions on A. All the coefficients of the second term falls off the diagonal of A and are thus negative, while the corresponding components of x are non-negative since $k \notin J$. Hence the second term is non-positive, reaching a contradiction. This shows that $x \geq 0$. A similar reasoning implies that $x \leq 0$, and x must actually be equal to zero. Hence A is invertible. ■

Lemma 4 *Let $\mathcal{S} = (I, D, (T_i)_{i\in I}, p, (u_i)_{i\in I})$ be a social choice problem, let μ be an incentive efficient mechanism, let $(\lambda, \alpha) \in (\times_{i\in I}\mathbb{R}^{T_i}_{++}) \times (\times_{i\in I}\mathbb{R}^{T_i\times T_i}_{+})$ be a pair of vectors that satisfy the conditions of Lemma 2, let $j \in I$, let $\bar{t}_j \in T_j$, and let $W^\lambda = \sum_{i\in I}\sum_{t_i\in T_i} \lambda_i(t_i)U_i(\mu|t_i)$. Then it is possible to construct a decision $\hat{d}$ and utility functions $(\hat{u}_i)_{i\in I}$ defined on $(D \cup \{\hat{d}\}) \times T$ such that*

1. *$\hat{u}_i(d, t) = u_i(d, t)$, for each $d \in D$, each $i \in I$, and each $t \in T$;*
2. *μ satisfies the conditions of Lemma 2 for (λ, α) in $\hat{\mathcal{S}} = (I, D\cup\{\hat{d}\}, (T_i)_{i\in I}, p, (\hat{u}_i)_{i\in I})$;*
3. *$\hat{U}_j(\hat{d}|\bar{t}_j) = W^\lambda/\lambda_j(\bar{t}_j)$;*
4. *$\hat{U}_j(\hat{d}|t'_j) = 0$ if $t'_j \in T_j \setminus \{t_j\}$;*
5. *$\hat{U}_i(\hat{d}|t_i) = 0$, for all $i \in N \setminus \{j\}$ and all $t_i \in T_i$.*

Proof: Let $f : \mathbb{R}^{I\times T} \to \mathbb{R}^{I\times T}$ be the linear transformation that maps any profile of ex-post utilities to its associated virtual utilities:

$$(f(u))_i(t) = \frac{1}{p(t_i)}[(\lambda_i(t_i) + \sum_{t'_i} \alpha_i(t'_i|t_i))u_i(t) - \sum_{t'_i} \alpha_i(t_i|t'_i)u_i(t_{-i}, t'_i)],$$

for each $u \in \mathbb{R}^{I\times T}$, each $t \in T$, and each $i \in I$. Let $\hat{f} : \times_{i\in I}\mathbb{R}^{T_i}$ be an analogue transformation for interim utilities:

$$(f(\mathfrak{u}))_i(t_i) = \frac{1}{p(t_i)}[(\lambda_i(t_i) + \sum_{t'_i} \alpha_i(t'_i|t_i))\mathfrak{u}_i(t_i) - \sum_{t'_i} \alpha_i(t_i|t'_i)\mathfrak{u}_i(t'_i)],$$

for each $\mathfrak{u} \in \times_{i\in I}\mathbb{R}^{T_i}$, each $t_i \in T_i$, and each $i \in I$. Lemma 3 implies that both f and $\hat{f}$ are invertible.

Let

$$v(t) = \max_{d\in D} \sum_{i\in I} v_i^{(\lambda,\alpha)}(d, t),$$

for each $t \in T$. Consider then the interim profile of utilities $\mathfrak{u}$ defined as follows:

$$\mathfrak{u}_j(\bar{t}_j) = \frac{W^\lambda}{\lambda_j(\bar{t}_j)}$$

$$(\forall (i,t_i) \neq (j,\bar{t}_j)) : \mathfrak{u}_i(t_i) = 0,$$

and let $\mathfrak{v}$ be the associated profile of interim utilities, i.e. $\mathfrak{v} = \hat{f}(\mathfrak{u})$. Notice that

$$\sum_{i \in I} \sum_{t_i \in T_i} p(t_i)\mathfrak{v}_i(t_i) = \sum_{t_j \in T_j} p(t_j)(\hat{f}(\mathfrak{u}))_j(t_j) = W^\lambda = \sum_{t \in T} p(t)v(t).$$

Lemma 1 implies that there exists $x : T \to \mathbb{R}$ such that

1. $(\forall t \in T) : \sum_{i \in I} x_i(t) = v(t)$, and
2. $(\forall i \in I)(\forall t_i \in T_i) : \mathfrak{v}_i(t_i) = \sum_{t_{-i}} p(t_{-i}|t_i)x_i(t)$.

The conditions of the present lemma are then satisfied if one defines $\hat{d}$ and the utility functions such that $u_i(\hat{d}, t) = (f^{-1}(x))_i(t)$, for all $i \in I$ and $t \in T$. Condition 1 is trivial, while condition 2 follows from Lemma 3 and the fact that $\sum_{i \in I} x_i(t) = v(t)$, $\forall t \in T$. The three remaining conditions follow from simple computations:

$$\sum_{t_{-i} \in T_{-i}} p(t_{-i}|t_i)u_i(\hat{d}, t) = \sum_{t_{-i} \in T_{-i}} p(t_{-i}|t_i)(f^{-1}(x))_i(t) = (\hat{f}^{-1}(\mathfrak{v}))_i(t_i) = \mathfrak{u}_i(t_i),$$

for each $t_i \in T_i$ and each $i \in I$. ■

Lemma 5 *Let I be a finite set of individuals, let $(T_i)_{i \in I}$ be a collection of type sets, let $p \in \Delta(T)$ be a common prior with full support, let $\lambda \in \times_{i \in I} \mathbb{R}^{T_i}_{++}$, and $\mathfrak{w} \in \times_{i \in I} \mathbb{R}^{T_i}_{++}$. Then there exists a collective decision d_{i,t_i}, for each combination (i, t_i), and a utility function $u_i : \{d_{j,t_j} | j \in I, t_j \in T_j\} \times T \to \mathbb{R}$, for each $i \in I$, such that the following conditions hold true for each (i, t_i):*

1. *$U_i(d_{i,t_i}|t_i) = [\sum_{j \in I} \sum_{t_j \in T_j} \lambda_j(t_j)\mathfrak{w}_j(t_j)]/\lambda_i(t_i)$;*
2. *$U_i(d_{i,t_i}|t'_i) = 0$, for all $t'_i \in T_i \setminus \{\bar{t}_i\}$;*
3. *$U_j(d_{i,t_i}|t_j) = 0$, for all $j \in N \setminus \{i\}$ and all $t_j \in T_j$;*
4. *$\sum_{j \in I} \frac{\lambda_j(t'_j)}{p(t'_j)} u_j(d_{i,t_i}, t') \leq \sum_{j \in I} \frac{\lambda_j(t'_j)}{p(t'_j)} \mathfrak{w}_j(t_j)$, for all $t' \in T$*

<u>Proof</u>: Let

$$v(t) = \sum_{j \in I} \frac{\lambda_j(t_j)}{p(t_j)} \mathfrak{w}_j(t_j),$$

for each $t \in T$. Notice that

$$\sum_{t\in T} p(t)v(t) = \sum_{t\in T}\sum_{j\in I} p(t_{-j}|t_j)\lambda_j(t_j)\mathfrak{w}_j(t_j) = \sum_{j\in I}\sum_{t_j\in T_j} \lambda_j(t_j)\mathfrak{w}_j(t_j).$$

Fix $i \in I$ and $\bar{t}_i \in T_i$. The previous computation implies that the conditions of Lemma 1 are satisfied for $\mathfrak{u} \in \times_{i\in I}\mathbb{R}^{T_i}$ defined as follows:

$$\mathfrak{u}_i(\bar{t}_i) = [\sum_{j\in I}\sum_{t_j\in T_j} \lambda_j(t_j)\mathfrak{w}_j(t_j)]/p(\bar{t}_i)$$

$$(\forall (j, t_j) \neq (i, \bar{t}_i)) : \mathfrak{u}_j(t_j) = 0.$$

Let $x : T \to \mathbb{R}$ be a function that satisfies the two conditions of Lemma 1. It is then easy to check that the four conditions of the present lemma are satisfied if one defines $d_{i.\bar{t}_i}$ and the utility functions so that $u_j(d_{i.\bar{t}_i}, t) = p(t_j)x_j(t)/\lambda_j(t_j)$, for all $t_j \in T_j$ and all $j \in I$. ■